HAWAIIAN PLANTERS' RECORD

Vol. LV 1956 No. 2

HAWAIIAN PLANTERS' RECORD

Vol. LV | 1956 | No. 2

A publication devoted to the sugar interests of Hawaii and issued by the Experiment Station for circulation among plantatons of the Hawaiian Sugar Planters' Association.

On the Cover: Aerial view of testing areas.

HAWAIIAN SUGAR PLANTERS' ASSOCIATION FOR 1956

OFFICERS

G. W. Sumner . . . . *President*
A. G. Budge . . . . *1st Vice President*
J. E. Russell . . . . *2nd Vice President*
Slator M. Miller . . . . *Vice President*
S. L. Platt . . . . *Secretary*
T. M. Brown . . . . *Treasurer*

EXECUTIVE COMMITTEE

A. G. Budge
J. E. Russell
G. W. Sumner
R. G. Bell
Boyd MacNaughton

EXPERIMENT STATION COMMITTEE

W. W. G. Moir, *Chairman*
F. W. Broadbent
W. M. Bush
R. A. Cooke, Jr.
F. A. Schaefer III

Associate Members

J. W. Anderson
John Cushnie
A. D. Ednie
A. J. Hebert
W. M. Moragne
F. E. Simpich, Jr.
H. A. Walker, Jr.
A. R. Grammer, *Secretary*

ADVISORY COMMITTEES OF THE STATION COMMITTEE

Field Engineering

H. A. Walker, Jr., *Chairman*
C. H. Ahlf
B. B. Alexander
A. J. Hebert
Malcolm MacNaughton
R. M. Frazier
R. C. Williamson
K. H. Berg
E. B. Holroyde
L. D. Baver

Factory Engineering

A. J. Hebert, *Chairman*
Wm. Bomke
W. J. Hull
W. Kenda
R. G. Watt
A. T. Hossack
L. M. Van Dreser
G. W. Tompkin
J. N. Orrick
L. D. Baver

By-Products

F. E. Simpich, Jr., *Chairman*
A. J. Hebert
W. Kenda
W. W. G. Moir
F. A. Schaefer III
C. E. S. Burns, Jr.
Alan E. Faye
F. C. Churchill
R. Johnston
L. D. Baver

SUGAR LOSSES—FIELD AND FACTORY

ROBERT H. HUGHES*

The Hawaiian sugar industry has led the world for many years in the intensity, efficiency and technology of its sugar production. Certainly Hawaii's position as the leader will be difficult to wrest from it, if and when that time is ever encountered. Tribute should be paid to the leaders throughout all branches of the industry for the outstanding achievements which have been attained in the past. There is no sign of any slackening in the pace of progress, and in this respect, optimism for the future can be enjoyed.

Despite the technological and productive efforts of all of us in sugar, however, an ominous cloud continually overhangs our efforts: that of not being able to produce our product at a price that will return a fair and just profit. This situation confronts us all with a constant challenge to increase our efforts toward production of our sugar at a lower cost.

One of the areas which is receiving an increasing amount of attention in our search for ways and means to produce at a lower cost, is that concerned with sugar losses—all sugar losses which occur between the time our maximum sugar yield is in the field of ripe cane to the time the finished product is delivered to storage at port terminals. The developing interest in this problem is an encouraging sign that work in this field will receive added impetus in the years ahead.

HISTORICAL BACKGROUND

Not long after the turn of the century, a tremendous interest developed in the problem of improving the efficiency of the milling of sugar cane. S. S. Peck, Chemist, and Dr. R. S. Norris, Assistant Chemist and later Sugar Technologist of the Experiment Station, HSPA, were notable leaders in this work which extended to virtually all of the Hawaiian mills and their operating personnel. During a relatively short period of time, the research on milling bore fruit in record efficiencies in many Hawaiian mills. Expansion of milling trains, development of improved knives and cane shredders, increased pressures on the mills, etc., all served to accomplish a notable increase in the recovery of sugar from the cane.

Some years later there followed intensive investigation into ways and means to reduce processing losses, particularly the losses of sugar in molasses. A method was developed to measure the efficiency of the recovery of sugar from molasses. The recognition of the effect of physical factors on massecuites and the develop-

* Mill Department Head of the Hawaiian Commercial and Sugar Company, Ltd., Puunene, Maui. This paper was first delivered at the HST meeting in November 1954, and was repeated at the annual meeting of the HSPA in December.

ment of revised sugar boiling techniques, as well as new designs of equipment, brought recoveries of sugar in processing to high levels and losses to low levels.

During the periods when losses were being studied in the factories, losses which could occur in the fields were also receiving close attention. Studies on the effects of burning cane and on the deterioration of cane because of delayed harvesting were carried on to successful conclusions in the development of procedures designed to keep these losses to absolute minimums. These procedures when applied with reasonable care and attention kept field losses small in comparison with the losses occurring in the factories.

The advent of mechanical harvesting and the attendant problems, the problems of producing sugar during the war years, and the problems connected with the immediate post-war adjustments, have demanded attention to such a degree that relatively little attention has been devoted to sugar losses until quite recently. Once more, however, interest is developing in the problem.

TYPES OF LOSSES

Relatively few losses occurring in the operations required to produce sugar are adequately measured, and consequently may not be properly appreciated. Notable work has been conducted by the Experiment Station during the last two years to measure some of those losses which heretofore have gone unmeasured. There remains to be devised, however, a method which will continuously indicate the extent of total sugar losses in a manner similar to the Recovery and Losses calculations in present-day mill control.

The sugar losses discussed in this report have been limited to those which occur between the time the mature crop is ready for harvest and the completion of processing. Other losses which are not included in this discussion are those which occur in the growing of the crop, such as those due to insect damage, disease damage, drought, wind damage, overripening, etc. The limited discussion presumes that these losses are adequately controlled by entomological, agricultural or other scientific means.

Sugar losses under consideration in this discussion fall into the following categories:

1. Burning
2. Deterioration
3. Harvesting and transporting
4. Cleaning
5. Milling
6. Processing

Definition of the categories of losses and an indication of the magnitude of the losses is important to further discussion.

1. Burning

There is considerable evidence which indicates that burning results in a loss of sucrose when the cane is subsequently treated in a washing process of some type. The magnitude of the loss caused by burning cane is not well defined by available data. However, the indicated loss due to burning and washing is 2.1 per cent of the sucrose in the field prior to burning.

2. Deterioration

Deterioration losses are those caused by bacterial and/or chemical actions which occur between the time the cane is killed by burning or harvesting and the time the cane is crushed. The magnitude of losses in this category will vary to a wide degree depending upon many factors, such as temperature, humidity, dampness, time, damage to the stalks, bacterial infection, etc. Generally speaking, plantation practice throughout the industry limits this loss by reducing to a minimum the time between burning and milling. The inclusion of this loss as part of the measurement of harvesting and transporting loss, reduces the need for its separate consideration.

3. Harvesting and Transporting

The losses which occur in this category fall into the following classes:

a) damage by equipment (harvesters, loaders, transport units)

b) cane left in the field following harvesting operations

c) cane left on the road

d) damage at the unloading station

The measurements which are used in this report limit the inclusion of losses to those caused by equipment damage and unloading at the mill. The losses caused by the cane left in the field and on the road are extensive, but are not considered in this study because data for representative areas throughout the industry are not available. Loss due to cane left in the field or on the road may account for as much as three per cent of the total sugar in the field at the time of harvest.

4. Cleaning

Sugar losses which occur in the handling of cane from the unloading station to the crushers of the mill, less the losses of sugar in cane included in refuse from the cleaner, are considered as cleaner losses. Meager data on the physical loss of cane in refuse are available but, conservatively, this loss is believed to be approximately 10 per cent of the total loss in the cleaner. The loss of sugar in the refuse from the cleaner is not incorporated in this study because of lack of data to relate the study on an industry basis.

5. Milling

Sucrose remaining in bagasse and sucrose destroyed by bacterial and/or chemical action during milling operations constitute the milling losses. Only the sucrose remaining in the bagasse can be adequately measured so that the loss in milling is presumed to be limited to the bagasse loss.

6. Processing

The processing losses are the most accurately measured of all those under consideration. Losses in processing occur in the form of physical and chemical losses in the boiling houses. Sucrose losses in filter cake and in drips from equipment represent physical losses, whereas bacterial and/or chemical losses, such as inversion, represent chemical losses in processing. Although accounted for in processing, sucrose remaining in final molasses is considered a loss.

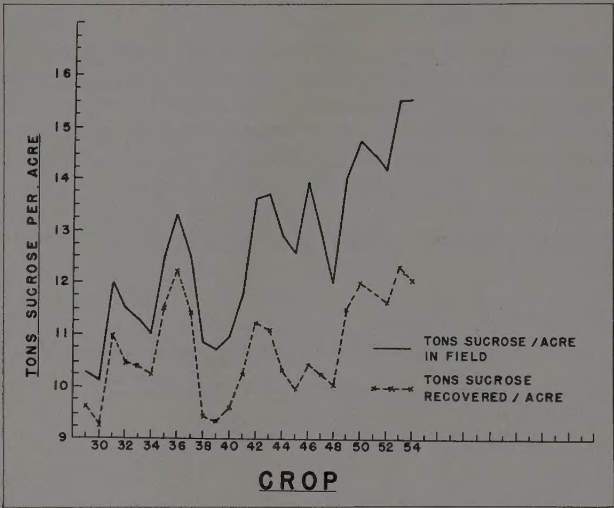

Figure 1

MAGNITUDE OF SUGAR LOSSES

The uncertainty of available data makes it necessary to limit the discussion to those types of losses which are most accurately measured and which will represent a conservative estimate of the magnitude of each loss. Consequently, losses which occur because of burning and deterioration are not considered in the remainder of this discussion, except insofar as they participate in harvest and transporting and cleaning categories. Losses due to cane left in the field, on the road, or in refuse from the cleaner, because of the inadequacy of data, are not incorporated in this report.

Data available from the study conducted by the Experiment Station staff on harvesting and cleaner losses, as well as mill control data, form the basis of the analyses incorporated in this report.

An analysis of the 25 crops from 1929 through 1953 at HC&SCo., Ltd., reveals the extent of the change in sugar losses which have occurred. This analysis is based upon the assumption that the losses due to burning and deterioration and not incorporated in harvest and transporting losses, are of the same magnitude today as they were 25 years ago.

The total losses which occur in processing today's crops are 350 per cent greater than those incurred 25 years ago, and represent an average loss of two tons of pure sucrose per acre. The change which has occurred during the 25 crops is graphically depicted in Figures 1, 2 and 3, as is also the trend in production and the recovery of pure sucrose.

Using for comparison the average sucrose per acre produced at HC&SCo. during the five crops at the beginning and the five crops at the end of the period, the change which has occurred is indicated in Table I.

Whereas the loss indicated for the one plantation may not be truly representative of the entire industry, it is reasonable to assume that the losses for the entire industry are somewhat proportionate, inasmuch as the factors which affected this

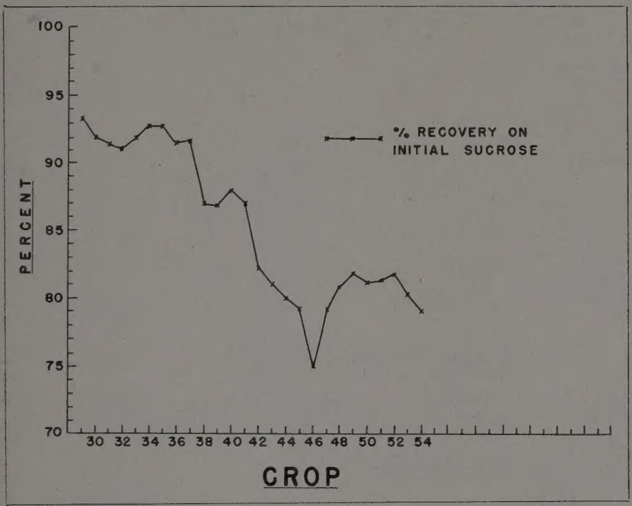

Figure 2

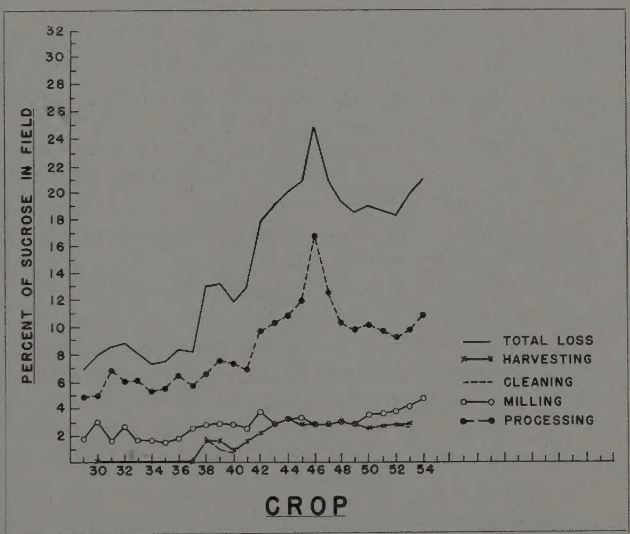

Figure 3

TABLE 1. TONS SUCROSE PER ACRE

Loss	1929-1933	1949-1953
Harvesting, Transporting	..	0.4
Cleaning	..	0.4
Milling	0.2	0.5
Processing	0.6	1.5
Total	0.8	2.8

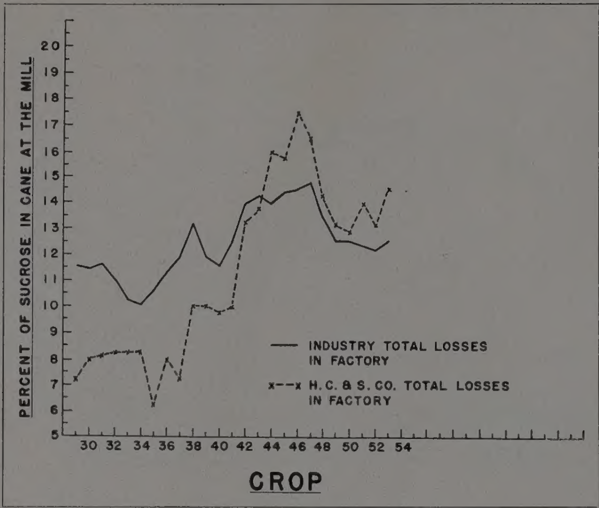

Figure 4

plantation have been experienced by virtually all other plantations at different times during the same period. Inasmuch as the losses in milling and processing constitute the major portion of the quantitative losses, a comparison of HC&SCo. and the industry in these categories serves to illustrate the general relationship existing between the two and supports the contention that the total losses for the sugar industry today are reasonably proportionate to the HC&SCo. losses. These data are presented in Figure 4.

The Hawaiian Crop would have been approximately 1,237,000 tons in place of 1,020,000 tons in 1952, and 1,321,000 tons in place of 1,099,000 tons in 1953, if the sucrose losses in those years had been at the level corresponding to that experienced in the 1929–1933 period.

It is reasonable to conclude that a problem of serious proportions exists when 15 to 25 per cent of the raw material in the field is not recovered in the final marketed product.

EVALUATION OF LOSSES

An analysis of the problem is incomplete without consideration of the economic losses involved. It is obvious that sugar which is lost in the field and in cleaning and milling is irretrievably lost, whereas some financial benefit accrues from any sugar included in molasses. The processing loss assumes different proportions when considered in the light of economics.

The following table presents estimates of the loss in revenue during the last five years and an estimated loss for 1954, based upon the returns received for sugar and molasses during the respective crop. The loss in revenue because of increased sucrose losses, using the 1929–1937 period as a standard, is indicated in one column, and the loss in revenue occurring because of the losses of sucrose in harvesting, transporting and cleaning under current factory recovery experience, is

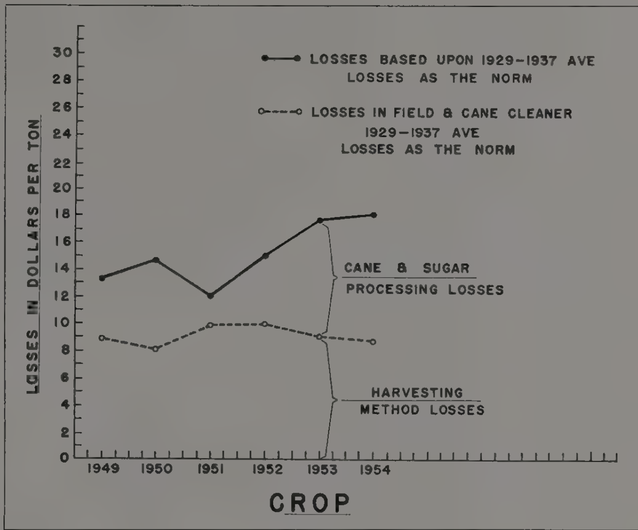

Figure 5

presented in the second column. The data are given in Table II and are graphically presented in Figure 5.

These data are indicative of conditions which existed at HC&SCo. Nevertheless it is believed that the estimates are representative of the losses throughout the industry where conditions of harvesting and cleaning are relatively equal. A similar analysis of data available for two other plantations indicates a loss in the area of \$9.00 per ton sugar for the 1953 crop. An increase in revenue to the Hawaiian Sugar Industry of \$10,000,000 annually is the potential available through an improvement in efficiency of the operations directly associated with present harvesting methods. An increase of \$20,000,000 in revenue annually is available when the industry can return to an efficiency in field and factory operations equivalent to that attained during 1929 to 1937.

Obviously, the potential increase in revenue will vary in accordance with the magnitude of physical losses and the returns for sugar and molasses.

In 1951, when molasses revenue was approximately \$39 per ton, the effect of greater or lesser amounts of sugar retained in molasses had a negligible effect upon income.

TABLE 2. LOSSES IN REVENUE FROM HARVESTING, TRANSPORTING AND CLEANING CANE

	1929-1937 Recovery Basis		Current Recovery Basis	
Year	\$/TS	\$000's	\$/TS	\$000's
1949	13.55	1885	8.92	1241
1950	14.84	2114	7.97	1136
1951	12.10	1832	9.92	1502
1952	15.08	1840	10.00	1485
1953	17.66	2842	9.25	1489
1954 (est.)	18.05	2803	8.60	1336

CAUSATIVE FACTORS IN THIS PROBLEM

Many factors can be enumerated that have a considerable effect upon sugar losses—field and factory. Several, however, have had outstanding influence, and a brief mention of the most outstanding is in order.

The greatest increase in sugar losses has occurred because of the harvesting methods adopted during the past 25 years. The need for changed harvesting procedures cannot be denied but a full appreciation of the ramifications of the harvesting method may have been overlooked. The harvesting method, itself a potent cause for sugar loss, required the construction of the cane cleaning plants, where large sugar losses also occur. The lack of cleaning plant efficiency allows foreign materials to damage the milling equipment and to increase amounts of sucrose in the bagasse, thus adding to existent losses. Similarly, increased soil in juices results in higher processing losses.

Losses in milling have increased markedly because of the higher grinding rates which have developed throughout the Hawaiian mills. The change from a 48-hour to a 40-hour work week has required the factories to increase the milling rate by approximately 20 per cent. Most factories have met the new grinding conditions without any changes in milling equipment. In some cases, the processing equipment has also been pressed into achieving a 20 per cent increase in rate without any increase in equipment.

The development of new cane varieties has increased yields to a marked degree. The newer varieties and the modified cultural practices they require have resulted in higher sugar losses, particularly with respect to processing losses. New varieties have also affected harvesting, transporting, cleaning and milling losses. Reports appearing in the proceedings of earlier Hawaiian Sugar Technologists' annual meetings bear witness to the effect of new varieties on the problem.

THE FUTURE

Much progress has been made in combatting field and factory sugar losses. Much remains to be accomplished. The goal for our efforts is a lucrative one.

The development of a harvesting method which will eliminate cane cleaners, and at the same time reduce harvesting and transporting losses, offers the greatest possibilities of economic benefit. Much progress has been made and a great store of knowledge has been accumulated in designing equipment and devising procedures to improve harvesting methods. Much remains to be done. Generally, the sugar losses have been of secondary consideration in developing new equipment for harvesting, but, be that as it may, all efforts in search of satisfactory harvesting equipment and procedures will contribute towards reduced sugar losses.

A determination to solve the problem of processing losses has been evident. Developments in de-ashing juices or syrups indicate that the factories may be on the threshold of a major advance in increasing the recovery of sugar and reducing the losses in molasses. The economic factors in the recovery of sugar from molasses weigh heavily in the practicability of new processes. Meanwhile, there can be no substitute for high quality work in present manufacturing procedures to keep sugar losses at a minimum.

Attention to milling losses has been limited in the past several decades. New attention must be focused upon them and a re-examination of the economic

factors may stir efforts in this direction. Several chemical products are presently offered as aids in reducing losses, but little data are available regarding the use of these aids under Hawaiian conditions.

In all of these fields, many hours of effort have been expended by the Experiment Station staff and by plantation and agency staff members toward the solution of the problem of reducing sugar losses. Progress has been made, but countless hours of work lie before us. The increase in the recovery of the sugar which is standing in the field and in which a major investment has been made, is of paramount importance. For each and every one engaged in the sugar industry, the record of the accomplishments thus far offers a challenge to carry on the study to a successful conclusion.

A METHOD FOR EVALUATING GENETIC PROGRESS IN A SUGAR CANE BREEDING PROGRAM

W. T. FEDERER[1]

The proposed method for evaluating the rate of genetic progress in the sugar cane breeding program of the Experiment Station, HSPA, is not considered as anything radically new but rather as a method for obtaining factual evidence relating to present and past procedures and ideas. A statistical procedure for a single test is discussed first. This is followed by a discussion of a procedure for analyzing several tests. The third section contains a statistical summarization of the results from four experiments selected by Dr. John N. Warner, Senior Geneticist, Expt. Sta., HSPA. The final section of the paper discusses the results and comments on the breeding program.

In the following, it will be assumed that the genetic material in the test is composed of two types, new strains (seedlings) and the check variety. It will also be assumed that there is sufficient seed of any seedling to plant only one plot. These assumptions are satisfied in the present 30′ x 30′ plantings. Here each seedling occupies a 30′ x 30′ plot and is placed adjacent to at least one plot of the check variety. The check variety occupies every third plot in the present scheme.

PROCEDURE FOR A SINGLE PLANTING

Consider that the k seedlings are planted in k plots with a number, n, of check plots scattered throughout the experimental area. The numbers n and k should be fairly large, say greater than 20–30, in order that the variance of a variance does not fluctuate too much. Theoretically, the n check plots should be randomly located in the experimental area in order to obtain an unbiased estimate of the environmental variance for the area under consideration. However, the estimated variance among systematically arranged check plots should not be too much larger than among randomly arranged check plots (Federer and Tanaka, 1955). Since the check plots arranged in a systematic manner, say every third plot, may be used to adjust seedling means for local variation, there may be a net gain obtained by arranging the check plots in a systematic fashion. (This assumes an appreciable correlation in yields between adjacent plots.)

The procedure follows:

Step (*i*) Compute the mean and variance for all check plots, thus:

$$\bar{y}_c = \left\{\sum_{i=1}^{n} Y_{ci}\right\}/n = Y_{c}./n \qquad (1)$$

[1] Principal Statistician, Experiment Station, HSPA, and Consultant in Statistics, PRI, while on sabbatical leave from Cornell University from September 1954 to September 1955.

and

$$s_c^2 = \left\{\sum_{i=1}^{n} Y_{ci}^2 - Y_{c\cdot}^2/n\right\}/(n-1) = \Sigma y_{ci}^2/(n-1), \tag{2}$$

where Y_{ci} = yield of check variety in i'th plot.

Step (*ii*) Compute the mean and variance for all seedling plots, thus:

$$\bar{y}_s = \left\{\sum_{j=1}^{k} Y_{sj}\right\}/k = Y_{s\cdot}/k, \tag{3}$$

and

$$s_s^2 = \left\{\sum_{j=1}^{k} Y_{sj}^2 - Y_{s\cdot}^2/k\right\}/(k-1) = \Sigma y_{sj}^2/(k-1), \tag{4}$$

where Y_{sj} = yield of j'th seedling.

Step (*iii*) Now, s_c^2 will be used as an estimate of the population variance σ_e^2 even though it is an overestimate (probably slight) if the check plots are arranged systematically. s_s^2 is an estimate of $\sigma_e^2 + \sigma_g^2$ where σ_e^2 is the true environmental variance and σ_g^2 is the true genetic variance in the population from which this random sample of k seedlings was selected. Therefore, the estimated genetic variance is

$$s_g^2 = s_s^2 - s_c^2. \tag{5}$$

That is, s_g^2 is the variance component which estimates σ_g^2, the true genetic variance.

Step (*iv*) If it is assumed that the true genetic merit of seedlings is normally and independently distributed with a general mean μ_s, estimated by $\bar{y}_s$, and a variance σ_g^2, estimated by s_g^2, a normal frequency curve for these data may be constructed using the sample values for μ_s and σ_g^2 (see Snedecor, 1946, Chapter 8).* This curve is given in Figure 1. The construction of the curve is instructive but not necessary since use may be made of tabled values of areas under the normal curve. (Although $\bar{y}_s$ and s_g^2 are random variates, useful information is derived from a group of such experiments.)

Step (*v*) From the location of μ_s and μ_c on the y axis in Figure 1, a partial evaluation of the seedlings relative to the check variety may be made. If $\bar{y}_s$ is considerably to the left of $\bar{y}_c$ relative to s_g, then the chances of obtaining a seedling better than the check variety are rather small. If, on the other hand, $\bar{y}_s$ is equal to or greater than $\bar{y}_c$ it is a relatively simple matter to select seedlings which are better than the check variety. If the breeding program has been in operation for a period of time, the latter situation is unlikely. It would be hoped that μ_s would not be over one or two σ_g units to the left of μ_c, and therefore genetic progress for higher yields would not be unduly difficult.

Step (*vi*) The seedling means are adjusted for the covariate adjacent check plot yields, in order to correct for yield gradients over the experimental area. The adjusted seedling yields are compared with $\bar{y}_c$, or perhaps some interval around $\bar{y}_c$, to determine which seedlings to select for future testing.

* If the distribution of true genetic merit is not too anormal, the method may still be used as an approximation.

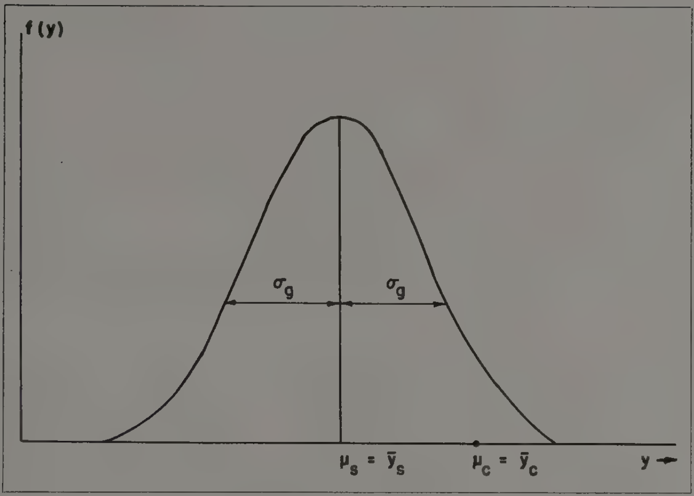

Figure 1. Normal curve constructed from the sample values $\bar{y}_s$ and s_g^2.

An alternative procedure to the above would be to make use of the normal curve using $\bar{y}_s = \mu_s$ and $s_s = \sigma_s = \sqrt{\sigma_e^2 + \sigma_g^2}$. This curve would have the same location parameter as in Figure 1 but the dispersion parameter, σ_s, would be larger. The *unadjusted* means are compared with $\bar{y}_c$, and all seedlings with means larger than $\bar{y}_c - as_e$, where a = a constant (e.g., -1, $-\frac{1}{2}$, $\frac{1}{2}$, 1, 2, etc.) is determined by the plant breeder, are selected for further testing. This procedure might be useful for fields with no established gradients.

PROCEDURE FOR SUMMARIZING DATA FROM SEVERAL TESTS

Suppose that a 30′ x 30′ test is located at each of p places and that the check variety with regard to yield is the same in each of these p places. Since the mean yield for the various areas may differ, some procedure other than the preceding one should be used for summarizing data from several fields. One such procedure that comes to mind is to use the differences of seedling means and the check variety mean for a given area; thus

$$Y_{shj} - \bar{y}_{ch} = d_{hj}, \tag{6}$$

where Y_{shj} = yield of j'th seedling at the h'th location and $\bar{y}_{ch}$ = mean yield of all check plots at the h'th location or place.

The variance of d_{hj} is

$$V(d_{hj}) = \sigma_e^2\{1 + 1/n_h\} + \sigma_g^2, \tag{7}$$

where n_h = number of check plots at the h'th place and where it is assumed that all estimates of genetic variance, s_{gh}^2, are estimates of the same parameter σ_g^2, that all estimates of environmental variances, s_{eh}^2, are estimates of the same parameter σ_e^2, that the correlation between Y_{shj} and $\bar{y}_{ch}$ is zero, and that there is no interaction of strains with fields or location.*

* If these assumptions are not valid, then it will be necessary to use a randomized and replicated design at each location in order to combine and interpret results over several locations.

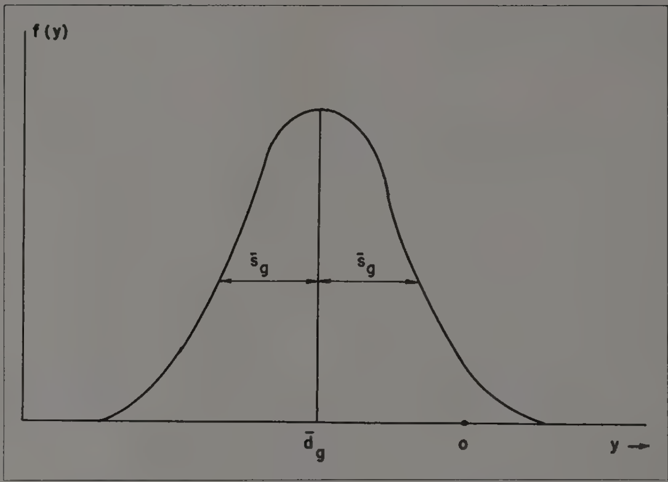

Figure 2. Normal curve constructed from the sample values $\bar{d}_g$ and $\bar{s}_g$.

Assuming normality, a normal curve could be constructed. The mean would now be

$$\bar{d}_g = \sum_{h=1}^{p} \left[\left\{ \sum_{j=1}^{k_h} d_{hj} \right\} / k_h \right] / p, \tag{8}$$

where k_h = number of seedlings harvested for yield at the h'th location. The variance used to construct the normal curve or to compute areas under the normal curve is

$$\bar{s}_g^2 = \left[\sum_{h=1}^{p} s_{gh}^2 (k_h - 1) \right] \Big/ \left[\sum_{h=1}^{p} (k_h - 1) \right]. \tag{9}$$

Graphically, the curve would be similar to that presented in Figure 2. The parameters of this curve would be taken equal to the sample values $\bar{d}_g$ and $\bar{s}_g$.

The differences d_{hj} would be compared with the point zero on the y axis. The selection procedure might be to save for further testing, seedlings which had values of $d_{hj} > 0$. A more rigid or a less rigid selection procedure might be adopted depending on the material in the experiment.

SUMMARIZATION OF YIELD RESULTS FROM FOUR WAIPIO EXPERIMENTS

(A2—Plant '54, B2—Plant '53, E—Plant '53, and D—1st Ratoon '51)

The data

The four 30′ x 30′ plantings selected for this study were grown on the Waipio Substation. Only a few of the seedlings were included in more than one planting. This means that there is only one replicate for most seedlings in the tests. The additional replication on the remaining seedlings was not taken into account.

The check plots were systematically arranged in such a fashion as to have a seedling plot adjacent to *at least one* check plot. Since there is only one replicate

for each seedling, it may safely be assumed that the arrangement of the seedling plots is a random one. If two or more replicates had been used, it would have been necessary to have a random arrangement of seedlings in the different replicates in order to validly use the statistical procedures discussed herein.

Variance and mean

The variance among check plots may be regarded as an estimate of the environmental variance of the experimental area. It is realized that this may be an overestimate of the environmental variance since the check plots are systematically spaced instead of being randomly arranged. However, for data of this type, the amount of overestimation is not considered serious (Federer and Tanaka, 1955).

The mean square for seedling plots is considered to be an estimate of the environmental variance plus the genetic variance. Therefore, the difference between the mean squares for the seedling and for the check plots yields an estimate of the genetic variance component in this material. Since the environmental variance may be somewhat overestimated, the genetic variance is somewhat underestimated. As stated above, the amount of overestimation, or underestimation, is not considered to be serious.

The various mean squares and genetic variances (formulae (2), (4), and (5)) for each of the four fields are given in Table 1 for tons of cane per acre (TCA). The same statistics for juice percentages (Y%C) and tons of sugar per acre (TSA) are given in Tables 2 and 3.

The various estimates of genetic variances for TCA are 103.02, 112.11, 50.14 and 51.69 with a weighted mean variance of 82.28. The corresponding genetic standard deviations are 10.15, 10.59, 7.08 and 7.19 with a weighted mean of 9.07. The genetic standard deviations are to be compared with the difference between the mean of the check plots and the mean of the seedlings. The corresponding differences of means for TCA are 3.0, 3.3, 5.1 and 12.1 with a weighted mean of

TABLE 1. STATISTICS ASSOCIATED WITH DATA FROM FOUR 30′ x 30′ SEEDLING TESTS FOR TONS-CANE-PER-ACRE (TCA)

Statistic	Test Number				Weighted[1] Mean
	A2	B2	E	D	
Mean of check plots	131.9	113.4	108.5	101.7	114.4
Mean of seedling plots	128.9	110.1	103.4	89.6	108.8
Mean square among check plots	142.88	140.98	126.90	81.29	123.47
Mean square among seedling plots	245.90	253.09	177.04	132.98	205.75
Genetic variance	103.02	112.11	50.14	51.69	82.28
Coefficient of variation for check plots	9.1	10.5	10.4	8.9	..
Regression of seedling plots on indices	.202	.306	.180	.641	.299
Correlation of indices and seedling plots	.131	.195	.136	.419	.199
Number of seedlings	64	64	46	59	233
Number of checks	32	32	24	30	118

[1] Means weighted by number of observations; mean squares weighted by degrees of freedom (formula (9)); the genetic variance weighted mean is obtained by subtraction, 205.75 − 123.47 = 82.28.

TABLE 2. STATISTICS ASSOCIATED WITH DATA FROM FOUR 30′ x 30′ SEEDLING TESTS FOR JUICE PERCENTAGES (Y%C = YIELD-PERCENT-CANE)

Statistic	Test Number				Weighted Mean
	A2	B2	E	D	
Mean of check plots..........	14.1	12.7	14.4	13.2	13.6
Mean of seedling plots..........	12.2	11.2	13.0	12.2	12.1
Mean square among check plots	0.78	0.68	0.51	0.22	0.56
Mean square among seedling plots....................	2.51	2.17	1.23	1.26	1.85
Genetic variance............	1.73	1.49	0.72	1.04	1.29
Coefficient of variation for check plots...............	6.2	6.5	4.9	3.6	..
Regression of seedling plots on indices................	.375	.416	.063	.354	.337
Correlation of indices and seedling plots.............	.157	.019	.031	.129	.147
Number of seedlings..........	64	64	46	59	233
Number of checks...........	32	32	24	30	118

5.6. Although the mean of the check variety is higher than the mean of the seedlings, the difference relative to the genetic standard deviation is not large enough to seriously impede selection for greater cane tonnage per acre.

With the mean values and the genetic standard deviations, one could, if desired, construct normal frequency curves for each field and for the combined results. However, for present purposes, a table of normal deviates is used to determine the percentage of times a normal deviate composed of $(\bar{y}_c - \bar{y}_s)/s_g$ is exceeded in random sampling. Thus, for field A, 3.0/10.15 = 0.296; a normal deviate of 0.296 is exceeded [1 − (.5 + .12)] 100 = 38% of the time, or roughly 4 times out of 10, in random sampling from a population of normal deviates (see areas under the normal curve in Handbook of Chemistry and Physics; Snedecor, 1946, Ch. 8).

The situation with regard to Y%C is not so encouraging. The corresponding differences in Y%C between the mean of the check and the mean of the seedlings

TABLE 3. STATISTICS ASSOCIATED WITH DATA FROM FOUR 30′ x 30′ SEEDLING TESTS FOR TONS-SUGAR-PER-ACRE (TSA)

Statistic	Test Number				Weighted Mean
	A2	B2	E	D	
Mean of check plots..........	18.6	14.4	15.6	13.4	15.5
Mean of seedling plots........	15.7	12.4	13.5	10.9	13.2
Mean square among check plots	2.68	3.51	3.24	1.48	2.71
Mean square among seedling plots....................	9.86	6.64	4.56	2.34	6.03
Genetic variance............	7.18	3.13	1.32	0.86	3.32
Coefficient of variation for check plots............	8.8	13.1	11.5	9.1	..
Regression of seedling plots on indices................	−.040	.076	−.222	.118	−.018
Correlation of indices and seedling plots.............	−.173	.046	−.152	.076	−.010
Number of seedlings..........	64	64	46	59	233
Number of checks...........	32	32	24	30	118

are 1.9, 1.5, 1.4, 1.0, and the corresponding genetic standard deviations are 1.32, 1.22, 0.85, 1.02. Thus, the mean of the check is one or more genetic standard deviation units higher than the mean of the seedlings in every one of the four tests. However, the situation is not hopeless since no difference between means is larger than two genetic standard deviation units.

Although it appears relatively easy to select sugar cane seedlings which exceed the check in TCA and not too difficult to select sugar cane seedlings which have better juices, the real objective is to select a sugar cane seedling which is high in both TCA and Y%C. Therefore, the important table would appear to be Table 3. Here we see that the mean differences in genetic standard deviation units are $2.9/\sqrt{7.18} = 1.1$, 1.1, 1.8, and 2.7 with a weighted average difference of $2.3/\sqrt{3.32} = 1.3$ genetic standard deviation units. The situation relative to selection of superior lines does not appear to be too bad here, although the task of selecting superior lines for TSA is somewhat more difficult than for TCA. Normal deviates larger than 1.3 are obtained approximately 10 per cent of the time in random sampling.

Optimum number of seedlings and replicates

With the above estimates of genetic and environmental variances, it is possible to determine the optimum number of replicates and seedlings to make maximum genetic progress by selecting the highest yielding entry in each test instead of selecting one at random. If one of the seedlings were randomly selected in each experiment, then we should expect to make no genetic progress. If, on the other hand, we always select the highest yielder and if $\sigma_g > 0$, we can expect to make progress. Following the method described by Federer (1951) for TSA, we use the quantities $s_e^2 = 2.71$, $s_g^2 = 3.32$, and $s_g^2/s_e^2 = 1.2$ and substitute these values in the formula:

$$s_g^2 \bar{x}_k / \sqrt{s_g^2 + s_e^2/r} = 3.32 \bar{x}_k / \sqrt{3.32 + 2.71/r}, \qquad (10)$$

where $\bar{x}_k$ is the average value obtained from selecting the largest member in a sample of size k from a normal population (see Federer, 1951) and where r is the number of replicates. The above formula is simpler computationally if environmental standard deviation units, $\sqrt{2.71}$, are used. Formula (10) may be written as:

$$1.2\, \bar{x}_k / \sqrt{1.2 + 1/r} \text{ environmental standard deviation units.} \qquad (11)$$

TABLE 4. POSITIVE WEIGHTED STANDARD DEVIATION UNITS FOR VARYING NUMBERS OF SEEDLINGS AND REPLICATES WHEN THE HIGHEST YIELDING SEEDLING IS SELECTED IN CONTRAST TO A RANDOMLY SELECTED ONE FROM THE SAME POPULATION

Number of seedlings	Av. value of largest member	Positive environmental standard deviation units with number of replicates indicated			
		1	2	4	8
25	1.97	1.59	1.81	1.96	2.05
50	2.25	1.82	2.07	2.24	2.35
75	2.40	1.94	2.21	2.39	2.50
100	2.51	2.03	2.31	2.50	2.62
150	2.65	2.14	2.44	2.64	2.76
200	2.75	2.23	2.53	2.74	2.87

Using the above formula and values, the effect of varying the number of seedlings and replicates is illustrated in Table 4. Another use of this table is to consider that a fixed number of experimental plots will be planted, e.g. 200. The various methods of planting 200 plots would be:

No. of reps.	No. of seedlings	Genetic Advance (in standard deviation units)
1	200	2.23
2	100	2.31
4	50	2.24
8	25	2.05

From the values for genetic advance it will be seen that the best combination is 100 seedlings and 2 replicates and the poorest combination is 25 seedlings and 8 replicates.

Relative efficiency of check plot method

As stated by Yates (1936), Cochran (1940), and Federer (1954), there are many ways of using check plots to adjust seedling plot yields. Only one method of constructing fertility indices from systematically arranged check plots and only two methods of using the indices are discussed here. The fertility index used is the yield of the adjacent check plot, if only one check plot is adjacent to the seedling plot. If two or more check plots border a seedling plot, the mean of the check plots is used as the fertility index.

The two methods of using the indices are:

$$Y_i - b(X_i - \bar{x}) \tag{12}$$

and

$$Y_i - (X_i - \bar{x}), \tag{13}$$

where Y_i is the yield of the i'th seedling, X_i is the fertility index associated with the i'th seedling, $\bar{x}$ is the mean of all indices in the test, and b is the regression coefficient of seedling plot yields on fertility indices. The first method uses the actual b value found for the data whereas the second method sets the b value equal to unity.

The various b values and correlation coefficients given in Tables 1, 2, and 3 are quite low. Only the b and r values for TCA in field D are considered to be significantly different from zero. If these tests can be considered as a representative sample of tests of this nature, there appears to be little or no relationship between seedling yields and fertility indices; the use of check plots to adjust seedling plot yields was a waste of area and labor for these experiments.

The sum of squares among adjusted seedling means (Federer, 1955, Chapter XVI) is

$$\Sigma y_i^2 - 2b\Sigma y_i x_i + b^2\Sigma x_i^2, \tag{14}$$

which for TSA for the 4 tests is equal to:

A2: $621.08 - 2(-.040)(-4.64) + (-.040)^2(116.36) = 620.89$, with 62 d.f.
B2: $418.03 - 2(.076)(11.59) + (.076)^2(152.98) = 417.15$, with 62 d.f.
E: $205.11 - 2(-.222)(-21.31) + (-.222)^2(95.80) = 200.37$, with 44 d.f.
D: $135.71 - 2(.118)(6.68) + (.118)^2(56.49) = 134.92$, with 57 d.f.

The sums of squares among adjusted seedling yields assuming $b = 1$ are:

$$\begin{aligned} &\text{A2: } 621.08 - 2(-4.64) + 116.36 = 746.72, \text{ with 63 d.f.}\\ &\text{B2: } 418.03 - 2(11.59) + 152.98 = 547.83, \text{ with 63 d.f.}\\ &\text{E: } 205.11 - 2(-21.31) + 95.80 = 343.53, \text{ with 45 d.f.}\\ &\text{D: } 135.71 - 2(6.68) + 56.49 = 178.84, \text{ with 58 d.f.} \end{aligned}$$

In every case, the sums of squares among adjusted yields assuming $b = 1$ are larger than those among adjusted yields using the actual regression coefficient. In fact, these sums of squares are larger than those for the unadjusted seedling yields. The assumption that $b = 1$ has made these data *more variable* than they were before adjustment. The net effect of the assumption that $b = 1$ when it is actually near zero is to reduce the effective number of replicates. In other words, with no adjustment we have one replicate, whereas we have only a fraction of a replicate when b is set equal to unity.

Any other method of constructing a regression coefficient other than using the actual regression is likely to reduce the effective number of replicates. In the above case the use of $b = 1$ lowered the effective number of replicates to about three-quarters of a replicate.

A method of comparing the efficiency of check plots is to compute the estimated environmental mean square without covariance and with covariance using the actual regression coefficient and using $b = 1$, and to take into account the extra land and resources utilized by the check plots. Thus, for TSA, the mean square among checks in field A2 is 2.68. The mean square among checks adjusted for covariance (msa) using $b = -.040$ is $\{(1 - r^2)\Sigma y^2_{ci}\}/(n - 2) = [1 - (-.173)^2\}83.07/30 = 2.69$. The adjusted mean squares for fields B2, E and D are:

$$\begin{aligned} &(1 - .046^2)108.92/30 = 3.62, \text{ with 30 d.f.}\\ &\{1 - (-.152)^2\}74.61/22 = 3.31, \text{ with 22 d.f.}\\ &(1 - .076^2)43.03/28 = 1.53, \text{ with 28 d.f.} \end{aligned}$$

The mean squares ($msa1$) among checks adjusted for regression, assuming $b = 1$, are approximated by the following for tests $A2$, $B2$, E and D, respectively:

$$\begin{aligned} &[746.72 - \{621.08 - 64 \times 83.07/32\}]/63 = 4.63\\ &[547.83 - \{418.03 - 64 \times 108.92/32\}]/63 = 5.52\\ &[343.53 - \{205.11 - 46 \times 74.61/24\}]/45 = 6.25\\ &[178.84 - \{135.71 - 59 \times 43.03/30\}]/58 = 2.20 \end{aligned}$$

TABLE 5. EFFICIENCY OF CHECK PLOTS IN SEEDLING TESTS FOR TSA

Test no.	msc = Mean sq. among check plots	msa = Adj. mean square using actual regression	msa1 = Adjusted mean square using b = 1	$\frac{\text{msc}}{\text{msa}} \times \frac{\text{No. seedling plots}}{\text{Total no. plots}}$	$\frac{\text{msc}}{\text{msa1}} \times \frac{\text{No. seedling plots}}{\text{Total no. plots}}$
A2	2.68	2.69	4.63	.66	.39
B2	3.51	3.62	5.52	.65	.42
E	3.24	3.31	6.25	.64	.34
D	1.48	1.53	2.20	.64	.45

When the mean squares adjusted for the actual regression are compared with the original mean squares, columns 2 and 3 of Table 5, we see that the loss of one degree of freedom, resulting from estimating b, more than offset the small gain due to utilizing the fertility indices. This loss does not consider the additional land and labor required for the check plots. In order to consider this, the ratio of the two mean squares is multiplied by the factor (number of seedling plots) (total number of plots). These results are given in the fifth column of Table 5. Here we see that the inclusion of check plots is about two-thirds as efficient as when the check plots are *omitted entirely*.

If the regression coefficient is assumed equal to unity, then the experimenter really is using an inefficient procedure. In the first place, the degrees of freedom for the mean squares given in column 4 of Table 5 can only be approximated (Cochran, 1951). When land use and labor are considered (last column of Table 5), we see that the method assuming $b = 1$ is only two-fifths as efficient as the method with no checks. Without considering land use and labor, the method of assuming $b = 1$ is still only about three-fifths as efficient as the method of no checks.

SOME COMMENTS

A detailed discussion of the breeding program of the Genetics Department, Experiment Station, HSPA, is given by Mangelsdorf (1953) and Warner (1953). Briefly, the program consists of the "bunch planting nursery," the 5′ x 3′ plot, 5′ x 6′ plot, 10′ x 15′ plot, 30′ x 30′ plot, and replicated test stages in that order. Seedlings usually appear but once in a given test in each of the stages except the replicated test stage. Some seedlings, depending on amount of "seed," may appear in more than one single plot test. The selection procedure for the six stages varies. Selections, on the basis of observational data, from the first four stages are made at the end of the first year of growth. No yield results are obtained until the 30′ x 30′ single plot stage which is usually harvested at 24 months of age. The most common age of harvest for commercial sugar cane in Hawaii is 24 months.

Since no prior yields are obtained, the "survivors" which enter the 30′ x 30′ single plot stage are not all selectable strains, and a number will be eliminated because of poor growth or condition during the second year of growth. A suggestion here would be to increase the seed supply as rapidly as feasible under the present observational selection procedure, and to use 40′ x 40′ plots harvested at 24 months of age. Periodic observations on cane condition should be made during the second year.

Since there is a possibility that "one-year-canes" mature faster than "two-year-canes," the former might be selected oftener than the "two-year-canes." If so, this would mean that an unduly large number of "one-year-canes" would be included in the 50–100 seedlings in a 30′ x 30′ test. This would mean that the effective number of "competitors" or selectable seedlings is much lower than the total number of seedlings included in the test. Two-year yield results should be obtained as quickly as possible in the program and before some good slow-starting "two-year-canes" are discarded because of poor growth during the first year.

Genetic progress for increased yields will be accelerated if the effective number of "competitors" in each test is increased. This tends to increase the mean yield of all seedlings, $\bar{y}_s$, with the possibility of having $\bar{y}_s$ exceed $\bar{y}_c$, the mean of the

check variety. A further acceleration of genetic progress is possible by using replicated designs (Federer, 1955) for the seedlings or by using a hoonuiaku design (Federer, 1956) for seedlings with a limited seed supply. Also, acceleration of genetic progress for increased yields is possible by a change in the breeding program (see Mangelsdorf, 1953, and Warner, 1953), which would increase the genetic variance, σ_g^2. A rigid preliminary selection procedure would need to be maintained in order that $\bar{y}_s$ would not be decreased more than one or two genetic standard deviation units below $\bar{y}_c$.

The use of hoonuiaku designs and the elimination of check variety plots in the early stages would free additional land, funds, and personnel which could be used to advantage in laying out replicated tests at several locations. The check plots for the early stages could easily be eliminated since the surrounding land is planted to a commercial variety of sugar cane.

Dr. J. L. Lush* kindly read an original version of this paper and pointed out a number of interesting features concerning the use of formula (10) and the statistics involved. The ratio $s_g^2/s_e^2 = 1.2$ is unusually large if one considers results from animal breeding experiments. However, this ratio is essentially equal to that obtained by Federer (1951) from experiments on corn single crosses. Since most sugar cane seedlings represent hybrids between sugar cane strains, the two ratios could be expected to agree fairly well. As pointed out by Dr. Lush, there are two reasons for the discrepancy between plant and animal breeding experiments; these are:

(i) The estimate of genetic variance obtained in hybrid material of this nature is the sum of the variances due to dominance, to epistasis, additive genetic effects, and to interactions. The additive genetic variance should be the only portion included in the ratio s_g^2/s_e^2.** This would tend to decrease the ratio, and selection for increased yields is more difficult than indicated from the data in Tables 1–3. The effect of a reduced genetic variance on optimum number of replicates and seedlings, for a fixed number of plots, is that more replicates and fewer seedlings give greater genetic progress than the reverse situation. In other words, the need for replicated tests becomes more important for a smaller ratio of genetic variance to environmental variance (Federer, 1951).

(ii) In animal experiments, observations are on individual animals, whereas in plant experiments, an observation on yield often comes from a group of plants. For example, in corn, the unit of observation may be a 2 hill x 10 hill plot or roughly 60 plants, and in sugar cane the unit of observation may be a 40′ x 40′ plot or roughly 200 plants. Therefore, in animal experiments the variance among strains is

$$\sigma_i^2 + m\sigma_p^2 + mr\sigma_g^2, \qquad (15)$$

where σ_i^2 = variance among individual animals (plants), σ_p^2 = variance among pens (plots), σ_g^2 = genetic variance, m = number of animals (plants) per pen (plot), r = number of replicates. Therefore the ratio of genetic to environmental variance for animal experiments is usually of the form:

$$\sigma_g^2/(\sigma_i^2 + \sigma_p^2) = G/E. \qquad (16)$$

* Iowa State College, in personal conversations and communications.

** This is true for sexually propagated plants. However, for asexually propagated plants like sugar cane, advantage may be taken of any epistatic or dominant effect.

In plant experiments the ratio of genetic and environmental variances is usually of the form:

$$m\sigma_g^2/(\sigma_i^2 + m\sigma_P^2) = \sigma_g^2/(\sigma_i^2/m + \sigma_P^2) = G/E \tag{17}$$

For m large, the variance among individual plants is almost eliminated, and the ratio obtained is essentially σ_g^2/σ_p^2 when plot totals are used.*

Use of formula (10) assumes a normal distribution of genotypic yields. Dr. Lush and Dr. A. J. Mangelsdorf, Experiment Station, HSPA, have pointed out that the true situation probably involves a curve which is truncated on the right. In any event, the distribution of true yields is not a normal distribution since there is both a lower (zero) and an upper (unknown) limit on yields of sugar cane. However, the method is still considered useful as a first approximation. The use of formula (10) on non-normal data of this type, indicates greater genetic progress and potentials than are actually possible.

A surprising result (to the author, at least) from the summarization of the data from the four Waipio Substation experiments, is the gross inefficiency of the check plot method (also see Cochran, 1940, and Yates, 1936). The conclusion reached here is that the method of check plots should be discarded and use should be made of a hoonuiaku design (Federer, 1956) when yield results are obtained for the first time. The method of assuming a regression coefficient of unity is grossly inefficient. It is assumed in the above that the four experiments are representative of tests of this type.

It might be argued that the correlation coefficients obtained are biased downward relative to the correlation between adjacent plots. This is true since the yield of a seedling plot may be expressed as

$$Y_j = Z_j + G_j, \tag{18}$$

where Z_j = yield of plot if it had been planted to the check variety and G_j = genetic merit of the j'th seedling. If X_j = yield of associated check plots, then $r_{XG} = 0$, $r_{ZG} = 0$, and r_{ZX} is *not* equal to 0. (We are assuming population values here.) The correlation between the seedling plot and adjacent check plot yields is

$$r_{yx} = \Sigma xy/\sqrt{(\Sigma z^2 + \Sigma g^2)\Sigma x^2} \tag{19}$$

when what we want is

$$r_{xz} = (\Sigma xy = \Sigma xz)/\sqrt{(\Sigma z^2)(\Sigma x^2)} \tag{20}$$

It is interesting to note that the regression coefficients are unbiased; thus,

$$b_{yx} = \Sigma xy/\Sigma x^2 = \Sigma xz/\Sigma x^2 = b_{zx} \tag{21}$$

Also, if we are willing to assume that the sum of squares Σz^2 is approximated by Σx^2, the regression coefficient equals the correlation coefficient. Even with this assumption, the resulting correlation coefficients (present regression coefficients) would not be anything to become excited about.

---

* The selection procedure determines whether formula (16) or (17) is appropriate. Since most plant selections are on an individual plant basis, formula (16) should be used in preference to formula (17) which is applicable when selecting a group of individuals.

One further point of interest concerning the regression and correlation coefficients is the comparison of the coefficients for TCA and Y%C with those for TSA. The coefficients for both Y%C and TCA are positive while those for TSA are near zero with two of them negative. Since TCA $\times$ Y%C = TSA for a given plot, it would be expected that the coefficients for TSA would at least stay positive. There must be a negative correlation between TCA and Y%C. An investigation on sugar cane uniformity trial data should cast some light on this phenomenon.

ACKNOWLEDGMENTS

Discussions on the breeding program of the Genetics Department of the Experiment Station, HSPA, with J. L. Lush, A. J. Mangelsdorf, J. N. Warner, R. Urata, and F. C. Denison, were extremely useful in orienting the studies reported in this paper. Appreciation for their cooperation and ideas is hereby expressed. The aid of R. K. Tanaka with computations and interpretations is greatly appreciated.

LITERATURE CITED

1. Cochran, W. G. 1940. A survey of experimental design. USDA Mimeo.
2. Cochran, W. G. 1951. Testing a linear relation among variances. Biometrics 7:17–32.
3. Federer, W. T. 1951. Evaluation of variance components from a group of experiments with multiple classifications. Iowa Agri. Expt. Sta. Res. Bul. 380.
4. Federer, W. T. 1954. Variety testing—some comments on the design of experiments. Paper presented at HST meetings, Nov. 16, 1954.
5. Federer, W. T. 1955. Experimental design—theory and application. Macmillan, N. Y.
6. Federer, W. T. 1956. Augmented (or Hoonuiaku) designs. Hawaiian Planters' Record 56: No. 2.
7. Federer, W. T. and Tanaka, R. K. 1955. A statistical study of the N$\times$K interaction for yield characters in group test no. 11. HSPA Expt. Sta. Spec. Rel. 116B. August.
8. Mangelsdorf, A. J. 1953. Sugar cane breeding in Hawaii, Part II, 1921–1952. Hawaiian Planters' Record 54:101–137.
9. Snedecor, G. W. 1946. Statistical methods, 4th ed. Iowa State College Press, Ames, Iowa.
10. Warner, J. N. 1953. The evolution of a philosophy on sugar cane breeding in Hawaii. Hawaiian Planters' Record 54:139–162.
11. Yates, F. 1936. A new method of arranging variety trials involving a large number of varieties. Jour. Agric. Sci. 26:424–455.

AUGMENTED (OR HOONUIAKU) DESIGNS[1]

WALTER T. FEDERER[2]

INTRODUCTION

In the present breeding program of the Experiment Station, HSPA, the seedlings surviving preliminary selection on the basis of agronomic characters other than yield, are planted in 30′ x 30′ single plot trials with a check variety occupying every third plot [4,6,7]. Yields per acre of cane and of sugar are obtained for each seedling in the 30′ x 30′ plantings. Replicated plots for yield determinations on the individual seedlings are not possible because of the scarcity of "seed" cane and because of the large plots required. Promising seedlings from the 30′ x 30′ plantings are advanced to replicated variety yield trials which include one or more of the standard commercial sugar cane varieties. (The plot size for the replicated variety test is 40′ x 40′ or larger.) The use of a check variety in every third plot is an inefficient design [4]; no experimental error is available from the present 30′ x 30′ single plot tests for making comparisons among the seedlings; it is often desirable to compare seedlings with more than one check variety. Consequently, it would be desirable to have an experimental design which would overcome these difficulties and which could be designed and analyzed with a minimum of difficulty.

The first experimental designs considered were the chain block and linked block designs constructed by Youden *et al.* [1,5,8,9]. These designs were developed to meet the needs of physicists and chemists whose experimental results are relatively homogeneous, thus allowing reliable estimates with only one or two measurements (plots). On the other hand, field results are relatively heteroge-

[1] This paper is a result of a discussion among F. C. Denison, E. B. Holroyde, J. F. Morgan, Jr., A. J. Mangelsdorf, T. Nakayama, and the author at Kahuku Plantation on March 8, 1955. An augmented randomized complete block design was installed at Kahuku on April 11, 1955. Additional experiments with this design have been installed in 1955 on the Kahuku, Ewa, and Pioneer Mill Plantations.

When these designs were first developed they were named "chain block designs" and this is the title used in early manuscripts and lectures. However, since these designs are different from the ordinary chain block and linked block designs, a new name should be applied to them. Since a name should be descriptive and popular, a number of people were consulted. To them, the author wishes to express his thanks. Some of the people offering suggestions and some of the suggested names are: W. S. Conner—designs with tagalongs; J. W. Tukey—caboose, penthouse, piggy-back, supplemental; O. Kempthorne—augmented; W. G. Cochran—augmented; and associates at Cornell University—bonded, fettered, expanded, ranged, topped, adjoined, "69", "combonded," bolted, latched, entangled, raveled, conjoined, etc. Of all names considered, "augmented" is the most descriptive. Another appropriate name for these designs might be "hoonuiaku" (pronounced hō-ō-nōō-ē-ă-kōō) since the designs were developed in Hawaii.

[2] Principal Statistician, Experiment Station, HSPA, and Consultant in Statistics, PRI, while on sabbatical leave from Cornell University from September 1954 to September 1955.

neous, necessitating several measurements (plots) over time and space in order to obtain reliable estimates. At first it appeared that some sort of chain block or linked block design was indicated because of scarcity of material rather than homogeneity of material, but it was found that these designs were not appropriate for the type of experiment desired for the breeding program. Therefore, it was necessary to develop new designs which possessed the desired characteristics. These designs have been denoted as *augmented* or *hoonuiaku* designs.

The purpose of this paper is to present the design and analysis for three designs in this new class of experimental designs. Examples are used to illustrate the design and analysis for the *augmented randomized complete block* design and for the *augmented latin square* design, which are considered to be the most useful for sugar cane breeding experiments. Other augmented designs are discussed elsewhere [3]. Augmented designs make it possible to achieve the following objectives in sugar cane breeding experiments:

(i) combine seedling and variety tests and thereby eliminate one-third of the present seedling plots (The varieties in the variety test replace the check-variety.),

(ii) provide efficient designs and a measure of experimental error for seedlings, and

(iii) provide comparisons among seedlings, among varieties, and between members of the two groups.

The designs have a much wider range of application than in genetic experiments. They are useful in all fields where it is desired to combine screening experiments on new material and preliminary testing experiments on promising material.

AUGMENTED (OR HOONUIAKU) DESIGNS

The class of experimental designs discussed in the present section is different from the class known as chain block designs, in that the latter class has a common link to only one other block, whereas the former class has several common links. However, it may be more enlightening to consider an augmented design as a standard design plus additional treatments (the seedlings) in the blocks or cells of the design, rather than as a design which has multiple links. The error mean square from all augmented designs, with the varieties repeated b times and the seedlings in once, is obtained from the analysis of the data on the varieties only. Hence, from an analysis standpoint, the consideration of a standard design with additional entries in the blocks or cells is more appropriate than the consideration of multiple linking.

As will be evident from the following, an augmented design may be formed from any standard design simply by enlarging the number of experimental units in the complete block, the incomplete block, the cell, etc., in the basic standard design and inserting seedlings in the additional plots. The adjustments for the v_1 seedling totals depend upon the design used and their location in the design. Whether or not the totals of the entries repeated b times, the varieties, require adjustment depends upon the standard design used. For example, no adjustments are required in the augmented randomized complete block design, but adjustments are required in the *augmented triple lattice* design.

I. AUGMENTED COMPLETELY RANDOMIZED DESIGN

The first experimental design that comes to mind for v_b varieties repeated b times and for v_1 seedlings each in one plot is the completely randomized design [2]. (The subscript on v denotes the number of times an entry is replicated.) The total experimental area is divided up into $bv_b + v_1 = N$ plots and the varieties and seedlings are allotted to the plots completely at random. No blocking is made in the entire area. By chance, all variety plots could fall in one corner of the experimental area. The analysis of variance for this well-known experimental design is:

Source of Variation	Degrees of Freedom
Entries..............	$v_b + v_1 - 1$
Among varieties.....	$v_b - 1$
Among seedlings....	$v_1 - 1$
Varieties vs. seedlings	1
Within entries = error.	$N - v_b - v_1 = v_b(b - 1)$
Total................	$N - 1$

The variety and seedling means do not require adjustment. The standard error of a mean difference for two varieties is

$$\sqrt{\text{error mean square } (2/b)}. \tag{I–1}$$

The standard error of a difference between two seedling means is

$$\sqrt{2 \text{ (error mean square)}}. \tag{I–2}$$

The standard error of a difference between a variety and a seedling mean is

$$\sqrt{\text{error mean square } (1 + 1/b)}. \tag{I–3}$$

This design is not recommended for sugar cane or other field experiments because it is highly improbable that a uniform area of N plots could be found. In most cases, it will be possible to reduce the error mean square considerably by proper stratification of the experimental area.

II. AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN

An augmented design which holds considerable promise for sugar cane variety and seedling tests is the one in which the v_b varieties are included once in each of the b blocks (i.e., replicated b times) and v_1 seedlings are included once in one of the b blocks. The blocks contain the v_b varieties plus n_{1j} seedlings, or a total of $v_b + n_{1j} = N_j$ plots.

$$bv_b + \sum_{j=1}^{b} n_{1j} = bv_b + v_1 = \sum_{j=1}^{b} N_j = N. \tag{II–1}$$

The total number of replicates on the $v_b + v_1 = v$ entries is

$$\frac{bv_b + v_1}{v_b + v_1} = 1 + \frac{(b-1)v_b}{v}. \tag{II–2}$$

To illustrate the grouping of varieties and seedlings, consider the following design for $v_b = 4$ (A,B,C,D), $v_1 = 11$ (e, f,g,h,i,j,k,l,m,n,o), and $b = 4$:

Group 1	Group 2	Group 3	Group 4
A	*A*	*A*	*A*
B	*B*	*B*	*B*
C	*C*	*C*	*C*
D	*D*	*D*	*D*
e	*h*	*k*	*n*
f	*i*	*l*	*o*
g		*m*	

Thus, there will be seven plots in three of the blocks and six plots in the fourth block. As is apparent from the above, the design is completely flexible with regard to the values of v_b, b, and v_1, but the n_{1j} should be as nearly equal as possible subject to the heterogeneity in the experimental area.

Randomization

As indicated earlier, there are b blocks with N_j plots in the j'th block. The randomization procedure follows:

(i) Randomly allot the varieties to the $v_b + n_{11} = N_1$ plots in block 1. This process is continued for the remaining blocks using a new random allotment in *each* block.

(ii) Randomly allot the v_1 seedlings to the remaining plots.

In setting up the b blocks, care should be taken to include a relatively homogeneous area within each block. The purpose of the blocking is to remove as much as possible of the heterogeneity in the experimental area and to have relatively homogeneous plots within each block.

Analysis

The analysis of the results for this experiment differs from that for orthogonal designs like the randomized complete block and latin square designs. The variety means, $\bar{y}_{i\cdot}$, require no adjustment for blocks since all varieties appear in all blocks. The mean effect, m, in the experiment is estimated as:

$$m = \frac{1}{v_b + v_1}\left\{Y_{\cdot\cdot\cdot} - (b-1)\Sigma\bar{y}_{i\cdot} - \Sigma n_{1j}r_j\right\}. \tag{II–3}$$

If n_{1j} is a constant, then the term $\Sigma n_{1j}r_j$ equals zero in (II–3). The effect of the i'th variety is obtained as

$$t_{bi} = Y_{i\cdot}/b - m = \bar{y}_{i\cdot} - m. \tag{II–4}$$

The effect of the j'th block is obtained as:

$$r_j = \frac{1}{v_b}\left(Y_{\cdot j\cdot} - \Sigma\bar{y}_{i\cdot} - \sum_{g=1}^{n_{1j}} Y_{1jg}\right). \tag{II–5}$$

The adjusted yield for a seedling plot is estimated from the following equation:

$$m + t_{1jg} = Y'_{1jg} = Y_{1jg} - r_j. \tag{II–6}$$

The above totals are defined in Table 1.

TABLE 1. SYMBOLIC REPRESENTATION OF YIELDS FROM AN AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN

Variety	Replicate 1	2	3	...	b	totals	means
A	Y_{bA1}	Y_{bA2}	Y_{bA3}		Y_{bAb}	$Y_{A\cdot}$	$\bar{y}_{A\cdot}$
B	Y_{bB1}	Y_{bB2}	Y_{bB3}		Y_{bBb}	$Y_{B\cdot}$	$\bar{y}_{B\cdot}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
v_b	Y_{bv_b1}	Y_{bv_b2}	Y_{bv_b3}		Y_{bv_bb}	$Y_{v_b\cdot}$	$\bar{y}_{v_b\cdot}$
Seedlings	Y_{111}	Y_{121}	Y_{131}		Y_{1b1}	..	..
	⋮	⋮	⋮	⋮	⋮		
	$Y_{11_{n1}}$	$Y_{12_{n2}}$	$Y_{13_{n3}}$		$Y_{1b_{nb}}$	..	..
Replicate totals	$Y_{\cdot 1\cdot}$	$Y_{\cdot 2\cdot}$	$Y_{\cdot 3\cdot}$		$Y_{\cdot b\cdot}$	$Y_{\cdot\cdot\cdot}$	..

where Y_{bij}=yield of i'th variety in j'th replicate and Y_{1jg}=yield of g'th seedling in j'th replicate.

The analysis of variance for an augmented randomized complete block design is:

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Block (ignoring entry)	$b-1$	B	..
Entries = varieties and sdlgs. (eliminating blocks)	v_b+v_1-1	SS_{vs}	V_{vs}
Varieties	v_b-1	SS_v	V_v
Sdlgs. and varieties vs. sdlgs.	v_1	SS_s	V_s
Intrablock error	$(v_b-1)(b-1)$	SS_e	E_e
Total	$N-1$	SS_t	..

An alternative analysis for the above design is:

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Block (eliminating varieties and seedlings)	$b-1$	SS_b	E_b
Entry (ignoring block)	v_b+v_1-1	T	..
Intrablock error	$(v_b-1)(b-1)$	SS_e	E_e
Total	$N-1$	SS_t	..

The various sums of squares are computed as follows:

$$SS_t = \sum_{j=1}^{b}\sum_{i=1}^{v} Y_{bij}^2 + \sum_{j=1}^{b}\sum_{g=1}^{n_{1j}} Y_{1jg}^2 - Y_{\cdot\cdot\cdot}^2/N. \tag{II–7}$$

$$B = \Sigma(Y_{\cdot j\cdot}^2)/(v_b + n_{1j}) - Y_{\cdot\cdot\cdot}^2/N. \tag{II–8}$$

The sum of squares due to m, r_j, t_{bi}, and t_{1jg} minus sum of squares due to m^* and r_j^* (assuming each t_{bi} and $t_{1jg} = 0$) is

$$SS_{vs} = mY\ldots + \sum_{j=1}^{b} r_j Y_{.j.} + \sum_{i=1}^{v_b} t_{bi} Y_{i.}$$

$$+ \sum_{j=1}^{b} \sum_{g=1}^{n_{1j}} t_{1jg} Y_{1jg} - \Sigma(Y^2_{.j.})/(v_b + n_{1j}). \quad \text{(II–9)}$$

$$SS_v = \Sigma Y^2_{i.}/b - (\Sigma Y_{i.})^2/bv_b. \quad \text{(II–10)}$$

$$SS_s = SS_{vs} - SS_v. \quad \text{(II–11)}$$

$$SS_e = SS_t - SS_{vs} - B. \quad \text{(II–12)}$$

$$T = \Sigma Y^2_{i.}/b + \sum_j \sum_g Y^2_{1jg} - Y^2_{\ldots}/N. \quad \text{(II–13)}$$

$$SS_b = B - (T - SS_{vs}). \quad \text{(II–14)}$$

The estimated variance of a difference between two variety means is

$$V(\bar{y}_{1.} - \bar{y}_{2.}) = 2E_e/b, \quad \text{(II–15)}$$

where E = the error mean square.

The variance of a difference between two seedling means depends upon the two seedlings involved. If two seedlings, say 1 and 2, are in the same block, the variance is

$$V(Y_{1j1} - Y_{1j2}) = V(Y'_{1j1} - Y'_{1j2}) = 2E_e. \quad \text{(II–16)}$$

If the two seedlings do not appear together in the same block, the variance of the mean difference is

$$V(Y'_{112} - Y'_{121}) = 2E_e(1 + 1/v_b): \quad \text{(II–17)}$$

The variance of a difference between a variety and a seedling mean is

$$V(\bar{y}_{i.} - Y'_{1jg}) = E_e(1 + 1/b + 1/v_b + 1/bv_b). \quad \text{(II–18)}$$

An Example

To illustrate the analysis of an augmented randomized complete block design, a portion of the data from a uniformity (blank) trial on Field 78 at Pioneer Mill Sugar Company, 1931, was used to construct an example (Table 2). The 3 = b

TABLE 2. FIELD ARRANGEMENT OF YIELDS (TCA) FOR AN AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN. DATA FROM BLANK TEST—PIONEER MILL NO. 78-1931

	Block 1							Total
Variety or seedling	*l*	*C*	*D*	*g*	*A*	*B*	*k*	
Yield	74	78	78	70	83	77	75	535 = $Y_{.1.}$
	Block 2							Total
Variety or seedling	*D*	*B*	*A*	*C*	*e*	*i*	..	
Yield	91	81	79	81	79	78	..	489 = $Y_{.2.}$
	Block 3							Total
Variety or seedling	*h*	*C*	*A*	*f*	*D*	*B*	*j*	
Yield	96	87	92	89	81	79	82	606 = $Y_{.3.}$
								1630 = $Y_{\ldots}$

blocks are from three different level ditches, and the yield character is tons of sugar cane per acre (TCA). Four hypothetical varieties, A, B, C, and D, and eight hypothetical seedlings, e, f, g, h, i, j, k, and l are used. Blocks 1 and 3 contain $v_b + n_{11} = v_b + n_{13} = 4 + 3 = 7$ entries each, and block 2 contains $v_b + n_{12} = 4 + 2 = 6$ entries. The block and grand totals are given in Table 2, and the variety and seedling totals are given in Table 3. As stated above, the variety means do not require adjustment for block effects. Hence, the arithmetic or unadjusted means for the varieties are used.

TABLE 3. VARIETY AND SEEDLING TOTALS, EFFECTS, AND ADJUSTED MEANS

Variety	Total	Means	Effects
A	254	84.67	3.6042
B	237	79.00	−2.0625
C	246	82.00	0.9375
D	250	83.33	2.2708
Total	987	329.00	4.7500

$m = 81.0625$
$r_1 = -3.25$
$r_2 = 0.75$
$r_3 = 2.50$

Seedling	Unadj. Mean	Adjusted Mean	Effects, t_{1jg}
e	79	78.25	−2.8125
f	89	86.50	5.4375
g	70	73.25	−7.8125
h	96	93.50	12.4375
i	78	77.25	−3.8125
j	82	79.50	−1.5625
k	75	78.25	−2.8125
l	74	77.25	−3.8125
Total	643	643.75	−4.7500

The mean effect, m, and the block effects, r_j, are estimated as follows (see formulae (II–3) and (II–5)).

$$m = \frac{1}{4+8}\left[1630 - 2(329) - \tfrac{1}{4}\left\{3(-13) + 2(3) + 3(10)\right\}\right] = 81.0625.$$

$$r_1 = \tfrac{1}{4}\left\{535 - 329 - 74 - 70 - 75\right\} = \frac{-13}{4} = -3.25.$$

$$r_2 = \tfrac{3}{4} = 0.75$$

$$r_3 = \tfrac{10}{4} = 2.50.$$

The sum of the block effects should add to zero, thus $-3.25 + 0.75 + 2.50 =$ zero.

The variety effects are computed from formula (II–4). For example, the effect for variety A is $\bar{y}_{A\cdot} - m = 84.6667 - 81.0625 = 3.6042$. The remaining variety effects are similarly computed and are presented in Table 3.

The seedling effects are computed from formula (II–6). For example, the effect for seedling l is $Y_{11l} - m - r_1 = 74 - 81.0625 - (-3.25) = -3.8125$. The remaining seedling effects are presented in Table 3. As a partial check, the sum of the variety and of seedling effects should equal zero.

TABLE 4. ANALYSES OF VARIANCE

Source of Variation	d.f.	Sum of Squares	Mean Square
Blocks (ignoring treatment)	2	360.0714	180.04
Treatments (eliminating blocks)	11	285.0954	25.92
Varieties	3	52.9167	17.64
Seedlings and varieties vs. seedlings	8	232.1787	29.02
Error	6	161.8332	26.9722
Total	19	807.0000	....
Blocks (eliminating treatments)	2	69.5001	
Treatments (ignoring blocks)	11	575.6667	....
Varieties	3	52.9167	....
Seedlings	7	505.8750	....
Seedlings vs. varieties	1	16.8750	....
Error	6	161.8332	....

The computations for the sums of squares in the analysis of variance (Table 4) are presented below (see formulae (II–7) to (II–14)):

$$SS_t = 74^2 + 78^2 + \ldots + 79^2 + 82^2 - \frac{1630^2}{20}$$

$$= 133{,}652 - 132{,}845 = 807.$$

$$B = \frac{535^2 + 606^2}{7} + \frac{489^2}{6} - \frac{1630^2}{20}$$

$$= 133{,}205.0714 - 132{,}845 = 360.0714.$$

$$SS_{vs} = 81.0625(1630) + [(-3.25)(535) + .75(489)$$

$$+ 2.50(606)] + [3.6042(254) + \ldots + 2.2708(250)]$$

$$+ [79(-2.8125) + 89(5.4375) + \ldots + 74(-3.8125)]$$

$$- 133{,}205.0714 = 133{,}490.1668 - 133{,}205.0714$$

$$= 285.0954.$$

$$SS_v = \frac{254^2 + \ldots + 250^2}{3} - \frac{987^2}{12} = 52.9167.$$

$$SS_s = 285.0954 - 52.9167 = 232.1787.$$

$$SS_e = 807 - 360.0714 - 285.0954 = 161.8332.$$

$$T = \frac{254^2 + \ldots + 250^2}{3} + 74^2 + \ldots + 75^2$$

$$- \frac{1630^2}{20} = 575.6667.$$

$$SS_b = 360.0714 - (575.6667 - 285.0954) = 69.5001.$$

If it is desired to obtain only an estimate of the error variance and to use one of the multiple range tests [see 2, Chapter II], a simple analysis is available. This analysis consists of using only the data from the varieties (Table 5) and using the analysis for a randomized complete block design for an orthogonal array. The

TABLE 5. YIELD DATA FOR VARIETIES

Variety	Block			Total
	1	2	3	
A	83	79	92	254
B	77	81	79	237
C	78	81	87	246
D	78	91	81	250
Total	316	332	339	987

fact that SS_b is equal to the blocks sum of squares in Table 6 within rounding errors, is not a coincidence. These sums of squares must be equal. Likewise, the error sums of squares in Tables 4 and 6 must be equal within rounding errors. This feature is a property of the augmented designs described herein.

The various standard errors of a mean difference are computed as follows (formulae (II–15) to (II–18)):

Between two variety means

$$\sqrt{\tfrac{2}{3}(26.9722)} = 4.24.$$

Between two seedlings in the same block

$$\sqrt{2(26.9722)} = 7.34.$$

Between two seedlings not in the same block

$$\sqrt{2(26.9722)(1 + \tfrac{1}{4})} = 8.21.$$

Between a variety and a seedling

$$\sqrt{26.9722(1 + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{12})} = 6.70.$$

For Tukey's hsd test [see 2, Chapter II], the first three standard errors of a mean difference above are multiplied by q_{05} (for 6 d.f. and 12 treatments) divided by $\sqrt{2}$. For the last standard error, it is suggested that the q_{05} value be multiplied by $\sqrt{E_e/n'}$ where $1/n'$ = harmonic mean of the coefficients for the standard errors of a difference. Thus, $n' = 4/(1 + \frac{1}{3} + \frac{1}{4} + \frac{1}{12}) = \frac{12}{5}$. It is not known how appropriate this approximation is for means with unequal numbers of observations.

TABLE 6. ANALYSIS OF VARIANCE FOR DATA IN TABLE 5

Source of Variation	d.f.	S.S.	M.S.
Blocks	2	69.5000	34.75
Varieties	3	52.9167	17.64
Error	6	161.8333	26.9722
Total	11	284.2500	....

III. AUGMENTED LATIN SQUARE DESIGN

In this design, the $v_b = b$ varieties are arranged in such a way that each variety appears once in a row or once in a column. This makes use of the principles of the latin square design. For example, consider the following design with $v_b = b = 3$ varieties (A, B, and C) in $b = 3$ rows and columns with $v_1 = 16$ seedlings (1,2,...,16, where the numbers are randomly allotted to the seedlings):

Columns

	1			2			3		
Row 1.....	C	1	2	A	3	4	5	B	6
2.....	A	12	11	10	B	9	8	C	7
3.....		B	13	14	15	C	16	A	

The number of seedlings in a given row and column, n_{1hi}, need not be the same; this is illustrated in the above design. The analysis will be somewhat simpler if n_{1hi} is a constant. Also, the n_{1hi} should be made as nearly equal as possible subject to the heterogeneity in the experimental area. It should be noted that the rows (or columns) could be laid end to end if desired [see 2].

Randomization

The randomization scheme for the augmented latin square design follows:

(i) Select a random arrangement for the $v_b = b$ varieties in a $b \times b$ latin square design [see 2, Chapter VI].

(ii) There are $n_{1hi} + 1 = N_{hi}$ plots in the h'th row and i'th column. The variety falling in the h'th row and i'th column is randomly allocated to one of the N_{hi} plots.

(iii) Assign random numbers to the v_1 seedlings and then distribute these to the remaining plots.

Analysis

The analysis of the data for an augmented latin square design follows that for an augmented randomized complete block design fairly closely. The variety means do not require adjustment; the seedling means must be adjusted for a row effect as well as a column effect. These adjustments are described below and in the example.

The solution for the mean, row, column, variety, and seedling effects are, respectively:

$$m = \frac{1}{v_1 + b}\left\{Y_{\cdot\cdot\cdot\cdot} - \left(b - 1 - \frac{2v_1}{b}\right)\Sigma\bar{y}_{b\cdot\cdot j} - \frac{1}{b}\sum_i\sum_h n_{1hi}Y_{bh\cdot\cdot} - \frac{1}{b}\sum_h\sum_i n_{1hi}Y_{b\cdot i\cdot}\right\}, \qquad \text{(III–1)}$$

$$r_h = \frac{1}{b}\left\{Y_{bh\cdot\cdot} - \sum_j \bar{y}_{b\cdot\cdot j}\right\}, \qquad \text{(III–2)}$$

$$c_i = \frac{1}{b}\left\{Y_{b\cdot i\cdot} - \sum_j \bar{y}_{b\cdot\cdot j}\right\}, \qquad \text{(III–3)}$$

$$t_{bj} + m = Y_{b\cdot\cdot j}/b = \bar{y}_{b\cdot\cdot j}, \text{ and} \tag{III–4}$$

$$t_{1hik} + m = Y_{1hik} - r_h - c_i, \tag{III–5}$$

where $v_1 = \sum_i \sum_h n_{1hi}$, $Y_{bh\cdot\cdot}$ = total of varieties in h'th row, $Y_{b\cdot i\cdot}$ = total of varieties in i'th column, $Y_{\cdot\cdot\cdot\cdot}$ = total of all entries in the experiment, $\bar{y}_{b\cdot\cdot j}$ = mean j'th variety, $Y_{b\cdot\cdot j}$ = total for j'th variety, and Y_{1hik} = yield of k'th seedling in the h'th row and i'th column.

The following analysis of variance is performed on the varieties only:

Source of Variation	d.f.	Sum of Squares	Mean Square
Row	$b - 1$	SS_r	..
Column	$b - 1$	SS_c	..
Variety	$b - 1$	SS_v	..
Residual or error	$(b - 1)(b - 2)$	SS_e	E_e
Total	$b^2 - 1$	..	..

The analysis of variance for varieties and seedlings is:

Source of Variation	d.f.	Sum of Squares
Row (ignoring entry and column)	$b - 1$	R
Column (ignoring entry; eliminating row)	$b - 1$	C
Entry (eliminating row and column)	$v_1 + b - 1$	SS_{vs}
Variety	$b - 1$	SS_v
Seedling and seedling vs. variety	v_1	SS_s
Error	$(b - 1)(b - 2)$	SS_e
Total	$b^2 + v_1 - 1$	....

The sum of squares for SS_{vs} may be obtained from the following relation:

Sum of squares for varieties and seedlings (ignoring row and column effects)

$$- (R - SS_r + C - SS_c) = SS_{vs}.$$

The computation of the other sums of squares is illustrated in the example.

The variance of a difference between two variety means, say 1 and 2, is

$$V(\bar{y}_{b\cdot\cdot 1} - \bar{y}_{b\cdot\cdot 2}) = 2E_e/b \tag{III–6}$$

where the experimental error is E_e.

The variance of a difference between adjusted means of two seedlings which appear together in the h'th row and i'th column, say $1hi1$ and $1hi2$, is

$$V(Y'_{1hi1} - Y'_{1hi2}) = 2E_e. \tag{III–7}$$

The variance of a difference between the adjusted means of two seedlings which appear in the same column but different rows (or in the same row but different columns) is:

$$V(Y'_{11i1} - Y'_{12i2}) = 2E_e(1 + 1/b) \tag{III–8}$$

The variance of a difference between adjusted means of two seedlings appearing in different rows and different columns, say 1111 and 1222, is

$$V(Y'_{1111} - Y'_{1222}) = E_e(2 + 4/b) \tag{III–9}$$

The variance of a difference between a variety mean and an adjusted seedling mean, say variety j and seedling 111, is

$$V(\bar{y}_{b\cdot\cdot j} - Y'_{1111}) = E_e(1 + 3/b - 2/b^2). \tag{III–10}$$

An Example

A hypothetical numerical example was constructed to illustrate the analysis for an augmented latin square design. Given that $m = 10$, $r_1 = -1$, $r_2 = 0$, $r_3 = 1$, $c_1 = -3$, $c_2 = -1$, $c_3 = 4$, $t_A = -1$, $t_B = -2$, $t_C = -3$, $t_d = 0$, $t_e = 2$, and $t_f = 4$, the following example was constructed:

Yields For Augmented Latin Square Design

Rows	Columns 1		2	3		Totals: All Entries	Totals: Varieties Only
1		(A) 5	(B) 6	(C) 10	(d) 13	$34 = Y_{\cdot 1\cdot\cdot}$	$21 = Y_{b\cdot 1\cdot}$
2		(B) 5	(C) 6	(e) 16	(A) 13	$40 = Y_{\cdot 2\cdot\cdot}$	$24 = Y_{b\cdot 2\cdot}$
3	(C) 5	(f) 12	(A) 9	(B) 13		$39 = Y_{\cdot 3\cdot\cdot}$	$27 = Y_{b\cdot 3\cdot}$
Totals (all)	$27 = Y_{\cdot\cdot 1\cdot}$		$21 = Y_{\cdot\cdot 2\cdot}$	$65 = Y_{\cdot\cdot 3\cdot}$		$113 = Y_{\cdot\cdot\cdot\cdot}$	....
Totals (var)	$15 = Y_{b\cdot 1\cdot}$		$21 = Y_{b\cdot 2\cdot}$	$36 = Y_{b\cdot 3\cdot}$		$72 = Y_{b\cdot\cdot\cdot}$	....

The variety and seedling (unadjusted) totals and means are:

$$
\begin{array}{ll}
Y_{b\cdot\cdot A} = 27 & \bar{y}_{b\cdot\cdot A} = 9 \\
Y_{b\cdot\cdot B} = 24 & \bar{y}_{b\cdot\cdot B} = 8 \\
Y_{b\cdot\cdot C} = 21 & \bar{y}_{b\cdot\cdot C} = 7 \\
Y_{1\cdot\cdot d} = 13 & \ldots. \\
Y_{1\cdot\cdot e} = 16 & \ldots. \\
Y_{1\cdot\cdot f} = 12 & \ldots.
\end{array}
$$

From formulae (III–1) to (III–5) we note that

$$
\begin{aligned}
m = \frac{1}{3+3}\Bigg[113 &- \left\{3 - 1 - \frac{2(3)}{3}\right\}\left\{9 + 8 + 7\right\} \\
&- \tfrac{1}{3}\left\{(1)(21) + (1)(24) + (1)(27)\right\} \\
&- \tfrac{1}{3}\left\{(1)(15) + (0)(21) + 2(36)\right\}\Bigg] = 10,
\end{aligned}
$$

$$r_1 = \tfrac{1}{3}\{\,21 - 24\,\} = -1,$$

$$r_2 = \tfrac{1}{3}\{\,24 - 24\,\} = 0,$$

$$r_3 = \tfrac{1}{3}\{27 - 24\} = 1,$$

$$c_1 = \tfrac{1}{3}\{15 - 24\} = -3,$$

$$c_2 = \tfrac{1}{3}\{21 - 24\} = -1,$$

$$c_3 = \tfrac{1}{3}\{36 - 24\} = 4.$$

$$t_A = 9 - 10 = -1,$$

$$t_B = 8 - 10 = -2,$$

$$t_C = 7 - 10 = -3,$$

$$t_d = 13 - (-1) - (4) - 10 = 0,$$

$$t_e = 16 - (0) - (4) - 10 = 2, \text{ and}$$

$$t_f = 12 - (1) - (-3) - 10 = 4.$$

Thus, all effects agree with the original values. This is a necessary condition for the correctness of the equations.

Before computing the sum of squares, the estimates m', c_i' and r_k' computed from the normal equations for the mean, the rows, and the columns with each t_{ij} and each t_{kl} set equal to zero, are required [3]. Using the relations that the sum of the row effects equals zero and that the sum of the column effects equals zero, the following normal equations for the constructed example are obtained:

$$12m' + c_1' + 2c_3' = Y_{\cdots} = 113,$$

$$4(r_1' + m') + c_3' = 34,$$

$$4(r_2' + m') + c_3' = 40,$$

$$4(r_3' + m') + c_1' = 39,$$

$$4(c_1' + m') + r_3' = 27,$$

$$3(c_2' + m') = 21, \text{ and}$$

$$5(c_3' + m') - r_3' = 65.$$

Solutions for the unknowns in the above equations follow:

$$m' = 8.9918919,$$

$$r_1' = -1.4932432,$$

$$r_2' = 0.0067568,$$

$$r_3' = 1.4864865,$$

$$c_1' = -2.5135135,$$

$$c_2' = -1.9918919, \text{ and}$$

$$c_3' = 4.4054054.$$

The adjusted seedling means (totals) are (formula (III–5)):

$$\begin{aligned} Y'_{1\cdot\cdot d} &= t_{113d} + m = Y_{1\cdot\cdot d} - r_1 - c_3 \\ &= 0 + 10 = 13 - (-1) - 4 = 10, \\ Y'_{1\cdot\cdot e} &= 2 + 10 = 16 - (0) - 4 = 12, \text{ and} \\ Y'_{1\cdot\cdot f} &= 4 + 10 = 12 - 1 - (-3) = 14. \end{aligned} \tag{III–11}$$

If only an estimate of the error mean square is desired then the analysis of variance in the central portion of Table 7 is calculated.* This analysis is identical to that for a standard $k \times k$ latin square. The analysis would be completed with the calculation of the various error variances given by formulae (III–6) to (III–10). Then, one of the multiple range tests [see 2 Ch. II] would be used to complete the comparison of means.

TABLE 7. ANALYSES OF VARIANCE

All Yields

Source of Variation	d.f.	Sum of Squares	Mean Square
Row (ignoring col. and entry)	2	5.1667	2.5834
Column (ignoring entry, eliminating row)	2	121.7635	60.8818
Entry (elim. row and column)	5	43.9865	8.7973
Variety	2	6.0000	3.0000
Remainder	3	37.9865	12.6622
Error	2	0.0000	0.0000
Total	11	170.9167	....
Correction for mean	1	1064.0833	....
Total (uncorrected)	12	1235.0000	....
On Yields Of Varieties Only			
Row	2	6	3
Column	2	78	39
Variety	2	6	3
Error	2	0	0
Total	8	90	....
Correction for mean	1	576	....
On All Yields Assuming Zero Entry Effect			
Row (eliminating column)	2	16.7635	....
Column (ignoring row)	2	110.1667	....
Residual	4	43.9865	....
Row (ignoring column)	2	5.1667	....
Column (eliminating row)	2	121.7635	....
Residual	4	43.9865	....

In some situations the analysis of variance given in the top part of Table 7 might be desired. The various sums of squares are computed as follows:

* In the example, the error sum of squares turns out to be zero as it should since no allowance was made for error variation in constructing the example.

Total sum of squares corrected for the mean (11 *d.f.*):

$$5^2 + 5^2 + \ldots + 13^2 + 13^2 - \frac{113^2}{12} = 1235 - 1064.0833 = 170.9167.$$

Row sum of squares (*ignoring column and entry effect*) (2 *d.f.*):

$$R = \frac{34^2 + 40^2 + 39^2}{4} - \frac{113^2}{12} = 5.1667.$$

Error sum of squares (2 *d.f.*):

total sum of squares uncorrected

$$-\{ mY_{\cdot\cdot\cdot\cdot} + \Sigma r_h Y_{\cdot h \cdot\cdot} + \Sigma c_i Y_{\cdot\cdot i \cdot}$$

$$+ \Sigma t_{bj} Y_{b \cdot\cdot j} + \Sigma\Sigma\Sigma t_{1hik} Y_{1hik} \} = \qquad \text{(III–12)}$$

$$1235 - [10(113) + \{ -1(34) + 0(40) + 1(39) \} + \{ -3(27)$$

$$- 1(21) + 4(65) \} + \{ -1(27) - 2(24) - 3(21) \} =$$

$$+ \{ 0(13) + 2(16) + 4(12) \}]$$

$$= 1235 - \{ 1130 + 5 + 158 - 138 + 80 \}$$

$$= 0,$$

as it should for this example.

Entry (*eliminating row and column effects*) *sum of squares* (5 *d.f.*):

$$SS_{vs} = mY_{\cdot\cdot\cdot\cdot} + \Sigma r_h Y_{\cdot h \cdot\cdot} + \Sigma c_i Y_{\cdot\cdot i \cdot} + \Sigma t_{bj} Y_{b \cdot\cdot j}$$

$$+ \Sigma\Sigma\Sigma t_{1hik} Y_{1hik} - \{ m' Y_{\cdot\cdot\cdot\cdot} + \Sigma r_h' Y_{\cdot h \cdot\cdot} + \Sigma c_i' Y_{\cdot\cdot i \cdot} \} \qquad \text{(III–13)}$$

$$= 1235 - \{ 8.8918919(113) + [-1.4932432(34)$$

$$+ .0067568(40) + 1.4864865(39)] + [-2.5135135(27)$$

$$- 1.8918919(21) + 4.4054054(65)] \}$$

$$= 1235 - 1191.0135180 = 43.9864820$$

Column (*ignoring entry, eliminating row*) *sum of squares* (2 *d.f.*):

$$C = m' Y_{\cdot\cdot\cdot\cdot} + \Sigma r_h' Y_{\cdot h \cdot\cdot} + \Sigma c_i' Y_{\cdot\cdot i \cdot} - \Sigma(Y^2_{\cdot h \cdot\cdot})/(b + n_{1h\cdot}) \qquad \text{(III–14)}$$

(where $n_{1h\cdot} = \sum_i n_{1hi}$)

$$= 1191.0135180 - \left(\frac{34^2 + 40^2 + 39^2}{4} \right)$$

$$= 1191.0135180 - 1069.25 = 121.7635180.$$

As a partial check,

$$121.7635180 + 5.1666667 + 43.9864820 = 170.9166667$$

$$= ss \text{ due to } m, r_h, c_i, t_{bj},$$

$$\text{and } t_{1hik} - (Y^2_{\cdot}\ldots)/(b^2 + \Sigma\Sigma n_{1hi}) \qquad \text{(III–15)}$$

$$= 1235 - \frac{113^2}{12}.$$

Row (*ignoring entry, eliminating column effect*) *sum of squares* (2 *d.f.*):

$$m'Y\ldots. + \Sigma r_h'Y_{\cdot h}\ldots + \Sigma c_i'Y_{\cdot\cdot i\cdot} - \Sigma(Y^2_{\cdot\cdot i\cdot})/(b + n_{1\cdot i}) \qquad \text{(III–16)}$$

$$= 1191.0135180 - \left\{\frac{27^2}{4} + \frac{21^2}{3} + \frac{65^2}{5}\right\}$$

$$= 1191.0135180 - 1174.25 = 16.7635180.$$

As a partial check,

$$16.7635180 = 5.1666667 - (110.1666667 - 121.7635180).$$

Also, the following sum of squares is of interest as a partial check on the other calculations:

Row + *column* (*eliminating entry effect*) *sum of squares* (4 *d.f.*):

$$mY\ldots. + \Sigma r_hY_{\cdot h}\ldots + \Sigma c_iY_{\cdot\cdot i\cdot} + \Sigma t_{bj}Y_{b\cdot\cdot j} + \Sigma\Sigma\Sigma t_{1hik}Y_{1hik}$$

$$- \left\{\Sigma Y^2_{b\cdot\cdot j}/b + \Sigma\Sigma\Sigma Y^2_{1hik}\right\} \qquad \text{(III–17)}$$

$$= 1235 - 582 - 569 = 84,$$

which should be the row + column sum of squares from the analysis of variance on variety yields only. For the example, 6 + 78 = 84 as it should.

The number of figures to the right of the decimal is larger than warranted by the accuracy of most data. The excessive number was carried because it was desired to show certain relations among the sums of squares free of rounding errors.

DISCUSSION

Although these designs were developed primarily for comparing seedlings and varieties developed by the sugar cane, pineapple, etc., geneticists, they have much wider usage. The entomologist and pathologist may wish to test a large number of new chemical compounds each year. The first test on such material is usually to decide if the new material shows promise, and anything that is obviously bad is discarded. Usually, only one replicate of such material is used. Also, the agronomist or physiologist may wish to compare a large number of new chemicals as weed sprays. The chemist in the laboratory may desire an estimate of experimental error for his material; he may use one of the chain block or linked block designs proposed by Youden *et al.* [5,8,9] or one of the designs presented herein.

Augmented designs make possible the elimination of the check plots and the combination of variety tests with seedling tests in the present sugar cane breeding program. Thus, one-third of the plots in seedling trials are eliminated and com-

parisons are possible with more than one "check" variety. The use of larger "blocks" for variety tests tends to increase the standard error for a variety mean, but this may be offset with additional replicates in the variety test with a net saving in total plots and with more information from the combined experiment than from the separate seedling and variety tests.

Since the analysis for the more complex augmented designs has not been developed, the question may arise as to what to do with seedlings for which there is enough available material to plant two plots. The seedling could be included once in two different tests or as two seedlings in one test. If the latter, the mean of the two adjusted means would be used as the seedling mean. This procedure is not efficient in that no use is made of the extra degree of freedom which could be allocated to error. This lack of efficiency will only be serious if the number of degrees of freedom for error is relatively low, say less than 10 to 16, or if several seedlings are repeated twice in each experiment. An analysis could be developed for these situations.

The question of systematic placement of the v_b varieties arises from time to time. If the varieties are systematically spaced, say every third plot, and if the analysis prescribed herein is used, the resulting mean differences and the experimental error variances are biased. Since the systematic placement of varieties invalidates known analyses, and since stratification, e.g., Design III versus Design I, accomplishes more than systematic placement of varieties, there appears to be no argument for using a systematic arrangement of varieties or "check plots" in seedling tests.

The success of any experimental design depends upon the ability of the experimenter to stratify or block the area into blocks which are relatively homogeneous within each block. For a large number of designs, the block size is constant from block to block. For chain block, linked block, and augmented designs, the block size may vary, but the blocks should be set up to minimize the variation among plots within the blocks.

SUMMARY

In response to the need for more efficient designs for comparing seedlings in the early stages of a breeding program, a new class of experimental designs, augmented designs, was developed. The analyses for the augmented randomized complete block and the augmented latin square designs are given and illustrated with numerical examples. Basically, the designs presented herein have one set of treatments (the varieties) repeated b times and a second set of treatments (the seedlings) appearing only once. A standard design (except for additional plots in the block) and analysis are used for the varieties, while the augmented design analysis is used for the varieties and seedlings together.

The designs are also useful in the fields of entomology, pathology, chemistry, physiology, agronomy, and perhaps others, for combining screening experiments on new material and preliminary testing experiments on promising material.

LITERATURE CITED

1. Eisenhart, C. 1951. On the statistical analysis of linked-block experiments (abstract). Biometrics 7:125.
2. Federer, W. T. 1955. Experimental design—theory and application. Macmillan, N. Y.
3. Federer, W. T. 1956. Augmented (or hoonuiaku) designs. BU–74–M. Feb.
4. Federer, W. T. 1956. A method for evaluating genetic progress in a sugar cane breeding program. Hawaiian Planters' Record, Vol. 55, No. 2.
5. Mandel, J. 1954. Chain block designs with two-way elimination of heterogeneity. Biometrics 10:251–272.
6. Mangelsdorf, A. J. 1953. Sugar cane breeding in Hawaii—Part II, 1921–1952. Hawaiian Planters' Record 54:101–137.
7. Warner, J. N. 1953. The evolution of a philosophy on sugar cane breeding in Hawaii. Hawaiian Sugar Planters' Record 54:139–162.
8. Youden, W. J. 1951. Linked blocks: A new class of incomplete block designs (abstract). Biometrics 7:124.
9. Youden, W. J. and Conner, W. S. 1953. The chain block design. Biometrics 9:127–140.

EXPERIMENT STATION STAFF

ADMINISTRATION

L. D. Baver........................*Director*
H. L. Lyon.....................*Consultant*

BUSINESS OFFICE

A. R. Grammer........*Senior Office Manager*
F. Wm. Littlejohn..*Associate Office Manager*

PUBLICATIONS AND VISUAL AIDS

L. D. Baver..........................*Editor*
R. J. Leffingwell..............*Consultant*
Lydia C. Nickerson.........*Associate Editor*
J. T. Yamamoto..*Photographer and Illustrator*

LIBRARY

Charlotta M. Hoskins...*Associate Librarian*
Jean Dabagh...........*Assistant Librarian*

BUILDINGS AND GROUNDS

William Sa Ning...........*Superintendent*

ISLAND REPRESENTATIVES

Hawaii

O. H. Lyman.....*Senior Island Representative*
K. Kunimitsu..*Assistant Island Representative*
A. Sato.................*Substation Overseer*
H. Miyamoto...........*Substation Assistant*
H. Oba................*Substation Assistant*
Y. Sugawara...........*Substation Assistant*

Oahu

F. C. Denison....*Senior Island Representative*
B. K. Nishimoto..........*Substation Overseer*
J. Bumanglag..........*Substation Assistant*
L. Nakatsuka..........*Substation Assistant*
H. Sakai...............*Substation Assistant*

Kauai

D. S. Judd.......*Senior Island Representative*
Y. Inouye..............*Substation Overseer*
C. G. Suero............*Substation Assistant*
A. T. Togikawa.........*Substation Assistant*

AGRONOMY

R. P. Humbert.........*Principal Agronomist*
A. S. Ayres...............*Senior Agronomist*
R. B. Campbell..........*Senior Agronomist*
N. S. Hanson............*Senior Agronomist*
A. H. Cornelison.......*Associate Agronomist*
R. E. Doty............*Associate Agronomist*
A. C. Trouse, Jr.......*Associate Agronomist*
H. Hagihara..........*Assistant Agronomist*
T. A. Jones...........*Assistant Agronomist*
J. A. Silva............*Assistant Agronomist*
Y. Yamasaki..........*Assistant Agronomist*
H. Asato...............*Agronomy Assistant*
A. Y. Ching............*Agronomy Assistant*
Geraldine Kau.........*Agronomy Assistant*
G. T. Yamamoto........*Agronomy Assistant*
M. Yamamoto..........*Agronomy Assistant*

BOTANY

H. L. Lyon.......................*Botanist*
Colin Potter.......*Nursery Superintendent*

BY-PRODUCTS

G. W. Aljian..................*Coordinator*
T. J. Nelson......................*Consultant*
S. B. Knapp..*Senior Pulp and Paper Technol.*

CHEMISTRY

F. E. Hance.............*Principal Chemist*
Richard Boyen.........*Executive Assistant*
Q. H. Yuen.................*Senior Chemist*
M. Doi.................*Associate Chemist*
T. S. Jones...............*Assistant Chemist*
H. M. Lee...............*Assistant Chemist*
Mon Yet Lum............*Assistant Chemist*
George Uyehara..........*Assistant Chemist*
George Koizumi...........*RCM Technician*
Henry Mau.......................*Analyst*
C. T. Nakagawa.......*Laboratory Technician*
J. Nishibun..........*Laboratory Technician*
E. S. Yamamoto.......*Laboratory Technician*

ENTOMOLOGY

F. A. Bianchi...........*Senior Entomologist*
J. W. Beardsley, Jr....*Assistant Entomologist*
R. E. Warner.........*Assistant Entomologist*
J. S. Rosa............*Laboratory Technician*

EXPERIMENTAL STATISTICS

J. G. Darroch........*Principal Statistician*
R. K. Tanaka.........*Assistant Statistician*

FIELD ENGINEERING

R. A. Duncan...............*Chief Engineer*
E. R. Bolles.........*Senior Design Engineer*
W. P. Burton.............*Project Engineer*
W. E. Hart...............*Project Engineer*
C. B. Holtwick...........*Project Engineer*
E. P. Morgan.............*Project Engineer*
T. Adachi................*Design Engineer*
S. Kawamura..............*Design Engineer*
S. Matsuoka..............*Design Engineer*
E. M. Norum......*Assistant Project Engineer*
G. S. Palmerton...*Assistant Design Engineer*
J. H. Kim.................*Design Draftsman*
P. K. Okamura...........*Design Draftsman*
Y. Isa.....................*Shop Foreman*
C. K. Mahoe...*Welder-Mechanic Journeyman*
K. Togashi...........*Machinist Journeyman*
E. T. Torricer....*Assistant Welder-Mechanic*
Jose Gatiuan.*Assist. Machinist & Storekeeper*

GENETICS

A. J. Mangelsdorf.......*Principal Geneticist*
J. N. Warner.............*Senior Geneticist*
Samuel Price.*Assoc. Geneticist (Coop. USDA)*
R. Urata..............*Associate Geneticist*
A. Doi.................*Assistant Geneticist*
M. Tanaka.....*Supt. Kailua Field Laboratory*
R. Chi...................*Genetics Assistant*
R. Imamura..............*Genetics Assistant*
N. Kimura...............*Genetics Assistant*
A. M. Kubota............*Genetics Assistant*
K. Nishimoto............*Genetics Assistant*
H. Terayama.............*Genetics Assistant*

GEOLOGY

D. C. Cox.................*Senior Geologist*
T. Hayashi...........*Laboratory Technician*

INSTRUMENT SHOP

G. E. Sloane...................*Consultant*
Gordon Furmidge....*Instrument Technologist*
G. E. Lambert.......*Instrument Technician*

ORGANIC CHEMISTRY RESEARCH

H. W. Hilton......*Principal Organic Chemist*
T. Yamamoto......*Assistant Organic Chemist*
E. S. Kurokawa......*Laboratory Technician*

PATHOLOGY

J. P. Martin..........*Principal Pathologist*
C. A. Wismer............*Senior Pathologist*
H. Koike...............*Assistant Pathologist*
A. S. Miura...........*Laboratory Assistant*

PHYSIOLOGY AND BIOCHEMISTRY

G. O. Burr.*Princ. Physiologist and Biochemist*
H. W. Brodie...........*Senior Physiologist*
Constance E. Hartt...*Associate Physiologist*
H. P. Kortschak........*Associate Biochemist*
T. Tanimoto...........*Associate Biochemist*
David Takahashi.......*Assistant Biochemist*
Robert E. Coleman....*Assistant Physiologist (Coop, USDA)*
Ada Forbes...........*Assistant Physiologist*
Robert M. Nakamura.*Laboratory Technician*
K. M. Sadaoka........*Laboratory Technician*
M. E. Wigmore........*Laboratory Technician*

SUGAR TECHNOLOGY

J. H. Payne...........*Principal Technologist*
G. E. Sloane............*Senior Technologist*
R. K. Hamilton........*Associate Technologist*
Toshio Moritsugu.....*Associate Technologist*
L. J. Rhodes..........*Associate Technologist*
Chen-Chuan Tu.......*Associate Technologist*
G. F. B. Appel, Jr......*Assistant Technologist*
F. J. Dupin..................*Scale Inspector*
A. C. Fong...........*Laboratory Technician*
H. S. Iwata..........*Laboratory Technician*
W. T. Shiramizu.......*Laboratory Technician*
Ray Spellman.........*Laboratory Technician*

TRAINING OF PERSONNEL

W. P. Alexander......*Principal Coordinator*

November 15, 1956

'/ ' ,

TABLE OF CONTENTS

Page

Sugar Losses—Field and Factory . . . . *Robert H. Hughes* 167

A Method for Evaluating Genetic Progress in a Sugar Cane Breeding Program . . . . *W. T. Federer* 177

Augmented (or Hoonuiaku) Designs . . . . *W. T. Federer* 191